

# Multiple-Antenna Fading Coherent Channels with Arbitrary Inputs: Characterization and Optimization of the Reliable Information Transmission Rate

Miguel R. D. Rodrigues, *Member, IEEE*

## Abstract

We investigate the constrained capacity of multiple-antenna fading coherent channels, where the receiver knows the channel state but the transmitter knows only the channel distribution, driven by arbitrary equiprobable discrete inputs in a regime of high signal-to-noise ratio (snr). In particular, we capitalize on intersections between information theory and estimation theory to conceive expansions to the average minimum-mean squared error (MMSE) and the average mutual information, which leads to an expansion of the constrained capacity, that capture well their behavior in the asymptotic regime of high snr. We use the expansions to study the constrained capacity of various multiple-antenna fading coherent channels, including Rayleigh fading models, Ricean fading models and antenna-correlated models. The analysis unveils in detail the impact of the number of transmit and receive antennas, transmit and receive antenna correlation, line-of-sight components and the geometry of the signalling scheme on the reliable information transmission rate. We also use the expansions to design key system elements, such as power allocation and precoding schemes, as well as to design space-time signalling schemes for multiple-antenna fading coherent channels. Simulations results demonstrate that the expansions lead to very sharp designs.

## Index Terms

Capacity, Constrained Capacity, Mutual Information, MMSE, Error Probability, Multiple-Antenna Fading Channels, Rayleigh Fading, Ricean Fading

Miguel R. D. Rodrigues was with Instituto de Telecomunicações–Porto and Departamento de Ciência de Computadores, Universidade do Porto, Portugal. He is now with the Department of Electronic and Electrical Engineering, University College London, United Kingdom [e-mail: m.rodrigues@ucl.ac.uk]. The work of M. R. D. Rodrigues was supported by Fundação para a Ciência e a Tecnologia, Portugal through the research project PTDC/EEA-TEL/100854/2008. This paper was presented in part at the IEEE International Symposium on Information Theory 2011 and 2012.

## I. INTRODUCTION

In recent years, the surge of interest in multiple-antenna communications systems has been due to the realization that the capacity of the canonical independent and identically distributed (i.i.d.) Rayleigh fading coherent channel, where the receiver knows the exact channel state but the transmitter knows only the channel distribution, scales as [1], [2], [3]:

$$C(\text{snr}) = \min(n_t, n_r) \log(\text{snr}) + \mathcal{O}(1) \quad (1)$$

so that, at high signal-to-noise ratio (snr), the potential gain in reliable information transmission rate of a multiple-antenna over a single-antenna system grows (linearly) with the minimum of the number of transmit or receive antennas,  $\min(n_t, n_r)$ , a quantity naturally known as the multiplexing gain or the degrees of freedom.

Of particular relevance has also been the characterization of the capacity of multi-antenna communications systems that embody prominent channel features that go beyond the canonical model. The crux of the characterizations, which for analytical tractability have been usually pursued in the asymptotic regimes of low- and high-snr, are affine expansions of the capacity in terms of fundamental performance measures. At low-snr, in order to capture the tradeoff between rate, bandwidth and power, it is appropriate to expand the capacity as an affine function of  $\frac{E_b}{N_0}|_{\text{dB}}$  as follows [4], [5]:

$$C\left(\frac{E_b}{N_0}\right) = \mathcal{S}_0 \cdot \left(\frac{\frac{E_b}{N_0}|_{\text{dB}}}{3 \text{ dB}} - \frac{\frac{E_b}{N_0 \min}|_{\text{dB}}}{3 \text{ dB}}\right) + o\left(\frac{E_b}{N_0} - \frac{E_b}{N_0 \min}\right) \approx \mathcal{S}_0 \cdot \log_2\left(\frac{E_b}{N_0} / \frac{E_b}{N_0 \min}\right) \quad (2)$$

where  $\frac{E_b}{N_0}|_{\text{dB}}$  is the (transmitted) energy per information bit  $\frac{E_b}{N_0}$  in dB,  $\frac{E_b}{N_0 \min}|_{\text{dB}}$  is the minimum (transmitted) energy per information bit  $\frac{E_b}{N_0 \min}$  required for reliable communication in dB, and  $\mathcal{S}_0$  is the capacity slope therein in bit/s/Hz/(3 dB). At high-snr, the capacity is expanded as an affine function of  $\text{snr}|_{\text{dB}}$  as follows [6], [7]:

$$C(\text{snr}) = \mathcal{S}_\infty \left(\frac{\text{snr}|_{\text{dB}}}{3 \text{ dB}} - \mathcal{L}_\infty\right) + o(1) \quad (3)$$

where  $\text{snr}|_{\text{dB}}$  denotes snr in dB,  $\mathcal{S}_\infty$  denotes the high-snr slope in bits/s/Hz/(3 dB) and  $\mathcal{L}_\infty$  denotes a zero-order term or a power offset in 3-dB units with respect to a reference channel with the same high snr slope but with unfaded and orthogonal dimensions whose expansion intersects the origin at  $\text{snr}|_{\text{dB}} = 0$ . Since the low- and high-snr quantities,  $\mathcal{S}_0$  and  $\frac{E_b}{N_0 \min}$  as well as  $\mathcal{S}_\infty$  and  $\mathcal{L}_\infty$  are a function of the random channel matrix, recourse to the elegant theory of random matrices [8] has disclosed the influence of various factors such as the number of transmit and receive antennas, antenna correlation, antenna polarization, line-of-sight components, spatially colored noise, interference, or even key signal features, on the reliable information transmission rate [4], [5], [7], [9], [10], [11], [12], offering a more realistic view of the potential of multiple-antenna communications. In general, the channel capacity, by virtue of the well-known  $\log \det(\cdot)$  expression [3], depends ultimately on the distribution of the eigenvalues of certain random matrices linked to the random channel matrix, a distribution which is known in some settings and in some asymptotic regimes [8]. Consequently, other characterizations valid for general snr, rather than only asymptotic snr regimes, have also been pursued.

It is well known, assuming that the channel variation over time is stationary and ergodic, that the capacity of the multi-antenna fading coherent channel is achieved by using (complex) Gaussian inputs, i.e., by using codewords whose elements are drawn from a zero-mean circularly symmetric complex Gaussian distribution that satisfy a transmit power constraint. However, it is also very relevant both from the theoretical and perhaps more importantly the practical perspective to study the constrained capacity of multi-antenna fading coherent channels driven by arbitrary (discrete) inputs. This is due to the fact that practical constraints pertaining to the transmission and reception of information often dictate the use of discrete inputs, such as PSK or QAM constellations, *in lieu* of the ideal Gaussian ones. Such a study, in addition to a deeper understanding, could also offer concrete guidelines for signal and system design.

The characterization of the constrained capacity of systems driven by non-Gaussian inputs poses a myriad of challenges due to the absence of explicit and tractable mutual information expressions. A innovative approach towards the resolution of this class of problems was put forth in [13], by exploiting connections between key quantities in information theory and estimation theory [14], [15]. By drawing upon the I-MMSE identity [14], Lozano *et al.* [13] have studied optimal power allocation policies for parallel non-interfering Gaussian channels with arbitrary inputs. Pérez-Cruz *et al.* [16] and Parayó and Palomar [17], [18] have in turn studied optimal power allocation and precoding for (interfering) multiple-input multiple-output (MIMO) Gaussian channels with arbitrary inputs (see also [19], [20], [21], [22], [23]). Optimum power allocation for multiuser OFDM systems with arbitrary signal constellations, both in scenarios where the transmitter knows the fading channel state and in scenarios where the transmitter knows only the fading channel distribution, has been addressed using identical techniques in [24].

This paper pursues the characterization of the reliable information transmission rate of multi-antenna fading coherent channels with arbitrary (discrete) inputs, by capitalizing on the intersections between information theory and estimation theory. The coherent channel model, where the receiver is assumed to know the exact channel state but the transmitter is only assumed to know the channel distribution, is a particularly relevant one because the use of bandwidth limited feedback channels between the receiver and the transmitter only enables the transmission of statistical, rather than instantaneous, channel state information in (fast fading) mobile communications systems. Due to the difficulty in constructing a general analytical characterization valid for all snr, the analysis focus on the key asymptotic regime of high snr in order to shed important insight about the fundamental communication limits. It is important to note, though, that the characterization of the reliable rate in coherent fading channels, where the receiver knows the exact channel state but the transmitter knows only the channel distribution, is considerably more complex than in MIMO Gaussian channels, where the (fixed) channel matrix is known exactly to the receiver and the transmitter (e.g., see [16], [17] and [18]). The most important aspect relates to the fact that the reliable rate is defined by the average of the mutual information with respect to the channel matrix distribution, rather than the mutual information conditioned on the (fixed) channel matrix only, as in [16], [17] and [18]. Consequently, we use the I-MMSE identity as a platform to explore other key analysis techniques, most notably, the machinery of asymptotic analysis and expansions, that lead to the exposure of the behavior of the reliable rate of multi-antenna fading coherent channels driven by arbitrary (discrete) inputs in the asymptotic regime of high snr. This paves the

way to the characterization of the constrained capacity of key multi-antenna fading coherent channel models as well as the design of key system elements.

#### *A. Contributions*

This paper contains various original contributions, which include:

- Analytic characterization of the constrained capacity of multi-antenna fading coherent channels driven by arbitrary discrete inputs in the asymptotic regime of high snr. The contribution reveals that, whilst at low-snr the capacity of a multiple-antenna coherent channel driven by Gaussian inputs and the constrained capacity of a multiple-antenna coherent channel driven by (proper complex) non-Gaussian inputs admit a similar characterization [4], at high-snr the behavior of the reliable rate for systems driven by the capacity-achieving Gaussian inputs and for systems driven by arbitrary discrete inputs is radically different; the contribution also reveals that the asymptotic characterization of the constrained capacity depends on the distribution of certain quadratic forms in Gaussian random variables, rather than the distribution of the singular values or the eigenvalues of certain random matrices. This aspect facilitates considerably the characterization of the constrained capacity of multi-antenna fading coherent channels.
- Analysis of the constrained capacity of various multi-antenna fading coherent channel models, including Rayleigh fading models, Ricean fading models and antenna-correlated models. The contribution emphasizes the impact of the channel properties as well as the signal geometry on the constrained capacity of multi-antenna fading coherent channels
- Design of system elements for various multi-antenna fading coherent channel models. In particular, the contribution illustrates with some detail the design of power allocation and precoding procedures as well as the construction of space-time schemes for various models.

In addition, the analysis also unveils intimate connections between the asymptotic behavior of key performance measures, namely, the average (non-linear) minimum mean-squared error, the average mutual information and the average error probability, whose interest may transcend the domain of application.

#### *B. Organization*

This paper is organized as follows: Section II describes the multiple-antenna fading coherent channel model. Sections III concentrates on the construction of a high-snr expansion for the reliable information transmission rate – the constrained capacity – of multiple-antenna fading coherent channels driven by arbitrary discrete inputs. The application of the expansions in the characterization of the reliable rate of the multiple-antenna canonical i.i.d. Rayleigh fading coherent channel as well as other multiple-antenna coherent channels that incorporate a variety of features is considered in Sections IV and V, respectively. We study in detail the effect on the reliable rate of Rayleigh fading, Ricean fading, the number of transmit and receive antennas, antenna correlation, as well as the signal properties. In turn, the application of the expansions in the design of communication system elements is considered in Section VI. In particular, we study power allocation in a bank of parallel independent fading channels,

power allocation and precoding in multiple-antenna fading channels, and space-time signal design. Section VII summarizes the main conclusions and contributions of the paper.

### C. Notation

We use the notation: Boldface uppercase letters denote matrices ( $\mathbf{X}$ ), boldface lowercase letters denote column vectors ( $\mathbf{x}$ ), and italics denote scalars ( $x$ ); the context defines whether the quantities are deterministic or random. The symbols  $\mathbf{I}$ ,  $\mathbf{0}$  and  $\text{diag}(d_1, d_2, \dots, d_n)$  represent the identity matrix, the null matrix and a diagonal matrix with diagonal elements  $d_1, d_2, \dots, d_n$ , respectively. The symbol  $\mathbf{e}_k$  represents a unit vector where the  $k$ th entry is equal to one and the other entries are equal to zero. The operators  $\det(\cdot)$ ,  $\text{tr}(\cdot)$ ,  $\text{rank}(\cdot)$ ,  $\text{vec}(\cdot)$  and  $\|\cdot\|$  represent the determinant, trace, rank, vectorization and Frobenius norm of a matrix, respectively. The operators  $(\cdot)^T$  and  $(\cdot)^\dagger$  represent the transpose and the Hermitian transpose of a matrix, respectively. The operator  $\mathbb{E}\{\cdot\}$  represents the expectation operation.  $\mathcal{CN}(\mu, \Sigma)$  denotes a circularly symmetric complex Gaussian random vector with mean  $\mu$  and covariance  $\Sigma$ .  $\Re(\cdot)$  and  $\Im(\cdot)$  denote the real part and the imaginary part of a complex number.  $H(\cdot)$  represents the entropy of a random variable or a random vector,  $H(\cdot|\cdot)$  represents the entropy of a random variable or random vector given another random variable or random vector and  $I(\cdot; \cdot)$  represents the mutual information between two random variables or random vectors. We also use the asymptotic notation:  $f(x) = \mathcal{O}(g(x))$  as  $x \rightarrow x_0$  if  $\lim_{x \rightarrow x_0} \left| \frac{f(x)}{g(x)} \right| < \infty$  and  $f(x) = o(g(x))$  as  $x \rightarrow x_0$  if  $\lim_{x \rightarrow x_0} \left| \frac{f(x)}{g(x)} \right| = 0$ .

## II. SYSTEM MODEL

We consider a fading channel with  $n_t$  transmit and  $n_r$  receive antennas which can be modeled as follows:

$$\mathbf{y} = \sqrt{\text{snr}} \cdot \mathbf{H}\mathbf{x} + \mathbf{n} \quad (4)$$

for a single use of the channel, where  $\mathbf{y} \in \mathbb{C}^{n_r}$  represents the vector of complex receive symbols,  $\mathbf{x} \in \mathbb{C}^{n_t}$  represents the vector of complex transmit symbols,  $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}) \in \mathbb{C}^{n_r}$  is a random vector which represents the noise and  $\mathbf{H} \in \mathbb{C}^{n_r \times n_t}$  is a random matrix which represents unit-power random channel gains between the various receive and transmit antennas, so that  $\mathbb{E}\{\text{tr}(\mathbf{H}\mathbf{H}^\dagger)\} = n_t n_r$ . We take the input to conform to an equiprobable multi-dimensional constellation with cardinality  $M$ , i.e.,  $\mathbf{x} \in \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M\}$  and  $\Pr(\mathbf{x}_1) = \Pr(\mathbf{x}_2) = \dots = \Pr(\mathbf{x}_M) = \frac{1}{M}$ , with  $\Sigma_{\mathbf{x}} = \mathbb{E}\{\mathbf{x}\mathbf{x}^\dagger\} = \frac{1}{n_t} \cdot \mathbf{I}$ . We also take the input, the noise and the channel to be independent. Therefore, the signal-to-noise ratio per receive antenna is given by:

$$\text{SNR} = \text{snr} \cdot \frac{\mathbb{E}\{\text{tr}((\mathbf{H}\mathbf{x})(\mathbf{H}\mathbf{x})^\dagger)\}}{n_r} = \text{snr} \cdot \frac{\mathbb{E}\{\text{tr}(\mathbf{H}\mathbf{H}^\dagger)\}}{n_t n_r} = \text{snr} \quad (5)$$

We consider a channel matrix that incorporates Rayleigh fading, Ricean fading as well as transmit and receive antenna correlation in a separable correlation model [25], [26], [27], [28], given by:

$$\mathbf{H} = \sqrt{\frac{K}{K+1}} \cdot \mathbf{H}_0 + \sqrt{\frac{1}{K+1}} \cdot \Theta_{\mathbf{R}}^{\frac{1}{2}} \mathbf{H}_{\mathbf{w}} \Theta_{\mathbf{T}}^{\frac{1}{2}} \quad (6)$$

where  $\mathbf{H}_0 \in \mathbb{C}^{n_r \times n_t}$  is a  $n_r \times n_t$  deterministic matrix,  $\mathbf{H}_{\mathbf{w}} \in \mathbb{C}^{n_r \times n_t}$  is a  $n_r \times n_t$  canonical complex Gaussian random matrix with independent zero-mean and unit-variance circularly symmetric complex Gaussian random entries

and  $\Theta_{\mathbf{T}} \in \mathbb{C}^{n_t \times n_t}$  and  $\Theta_{\mathbf{R}} \in \mathbb{C}^{n_r \times n_r}$  are  $n_t \times n_t$  and  $n_r \times n_r$  unit-diagonal Hermitian positive semi-definite matrices with the correlation coefficients between the  $n_t$  transmit and the  $n_r$  receive antennas, respectively. The Ricean K-factor  $K$  corresponds to the ratio between the deterministic and random component energies. Since the term  $\mathbf{H}_0$  corresponds to line-of-sight or diffracted components,  $\mathbf{H}_0 = \mathbf{a}_{\mathbf{R}} \mathbf{a}_{\mathbf{T}}^\dagger$  where the vectors  $\mathbf{a}_{\mathbf{T}} \in \mathbb{C}^{n_t}$  and  $\mathbf{a}_{\mathbf{R}} \in \mathbb{C}^{n_r}$  are associated with the transmit and receive array responses to a plane wave so that  $\|\mathbf{a}_{\mathbf{T}}\|^2 = n_t$  and  $\|\mathbf{a}_{\mathbf{R}}\|^2 = n_r$ . Note that in channels with Rayleigh fading, where line-of-sight components are absent,  $K = 0$ , whereas in channels with Ricean fading, where line-of-sight components are present,  $K \neq 0$ ; in addition, in channel models without transmit or receive antenna correlation  $\Theta_{\mathbf{T}} = \mathbf{I}$  and  $\Theta_{\mathbf{R}} = \mathbf{I}$  and in channel models with transmit and receive antenna correlation  $\Theta_{\mathbf{T}} \neq \mathbf{I}$  and  $\Theta_{\mathbf{R}} \neq \mathbf{I}$ .

We assume that the receiver knows the channel matrix realization and that the transmitter knows only the channel matrix distribution. We also assume that the sequence of random channel matrices over time is stationary and ergodic. Consequently, the constrained capacity, achieved by coding over multiple fading blocks, is given by:

$$\mathbb{E}_{\mathbf{H}} \left\{ I(\mathbf{x}; \sqrt{\text{snr}} \cdot \mathbf{H}\mathbf{x} + \mathbf{n} | \mathbf{H}) \right\} \quad (7)$$

where  $I(\mathbf{x}; \sqrt{\text{snr}} \cdot \mathbf{H}\mathbf{x} + \mathbf{n} | \mathbf{H})$  is the mutual information between input vector  $\mathbf{x}$  and the output vector  $\mathbf{y} = \sqrt{\text{snr}} \cdot \mathbf{H}\mathbf{x} + \mathbf{n}$  conditioned on a realization of the channel matrix  $\mathbf{H}$ . The goal is to characterize, as well as optimize, the constrained capacity in the asymptotic regime of high snr.

### III. CHARACTERIZATION OF THE CONSTRAINED CAPACITY

We provide a characterization of the constrained capacity of multiple-antenna fading coherent channels with arbitrary equiprobable discrete inputs in the regime of high snr. This characterization, which is the crux of the study of the effect of common channel parameters and models on the system performance as well as the design of key system elements in subsequent sections, is based on an asymptotic expansion of the constrained capacity that portrays its behavior as a function of key system parameters in the regime of high snr.

The definition of the asymptotic behavior is based on the procedure: First, we consider lower and upper bounds to the MMSE associated with the estimation of the noiseless output given the noisy output of the channel model in (4), for a fixed channel matrix, given by:

$$\text{mmse}(\text{snr}; \mathbf{H}) = \mathbb{E} \left\{ \left\| \mathbf{H}\mathbf{x} - \mathbf{H} \mathbb{E} \left\{ \mathbf{x} | \sqrt{\text{snr}} \cdot \mathbf{H}\mathbf{x} + \mathbf{n} \right\} \right\|^2 | \mathbf{H} \right\} \quad (8)$$

Second, we consider lower and upper bounds to the average value of the MMSE associated with the estimation of the noiseless output given the noisy output of the channel model in (4), for a random channel matrix, given by:

$$\overline{\text{mmse}}(\text{snr}) = \mathbb{E}_{\mathbf{H}} \left\{ \text{mmse}(\text{snr}; \mathbf{H}) \right\} = \mathbb{E}_{\mathbf{H}} \left\{ \mathbb{E} \left\{ \left\| \mathbf{H}\mathbf{x} - \mathbf{H} \mathbb{E} \left\{ \mathbf{x} | \sqrt{\text{snr}} \cdot \mathbf{H}\mathbf{x} + \mathbf{n} \right\} \right\|^2 | \mathbf{H} \right\} \right\} \quad (9)$$

We then consider upper and lower bounds to the mutual information between the input and the output of the channel model in (4), for a fixed channel matrix, given by:

$$I(\text{snr}; \mathbf{H}) = I(\mathbf{x}; \sqrt{\text{snr}} \cdot \mathbf{H}\mathbf{x} + \mathbf{n} | \mathbf{H}) \quad (10)$$

as well as upper and lower bounds to the average value of the mutual information between the input and the output of the channel model in (4), for a random channel matrix, given by:

$$\bar{I}(\text{snr}) = \mathbb{E}_{\mathbf{H}} \{I(\text{snr}; \mathbf{H})\} = \mathbb{E}_{\mathbf{H}} \left\{ I(\mathbf{x}; \sqrt{\text{snr}} \cdot \mathbf{H}\mathbf{x} + \mathbf{n} | \mathbf{H}) \right\} \quad (11)$$

by capitalizing on the I-MMSE identity and counterparts [14], [15]. Finally, we expose the asymptotic behavior as  $\text{snr} \rightarrow \infty$  of the average value of the MMSE and the average value of the mutual information, which leads to the asymptotic behavior as  $\text{snr} \rightarrow \infty$  of the constrained capacity, by capitalizing on the machinery of asymptotic analysis and asymptotic expansions [29].

The bounds are expressed in terms of the squared Euclidean distance between the (noiseless) receive vectors  $\mathbf{H}\mathbf{x}_i$  and  $\mathbf{H}\mathbf{x}_j$  associated with the transmit vectors  $\mathbf{x}_i$  and  $\mathbf{x}_j$  given by:

$$d_{ij}^2(\mathbf{H}) = \|\mathbf{H}\mathbf{x}_i - \mathbf{H}\mathbf{x}_j\|^2 \quad (12)$$

as well as their probability density functions  $p_{d_{ij}^2}(d_{ij}^2)$  and the higher-order derivatives  $p_{d_{ij}^2}^{(n)}(d_{ij}^2) = d^n \left( p_{d_{ij}^2}(d_{ij}^2) \right) / d(d_{ij}^2)^n$ ,  $n = 1, 2, \dots$

Let us consider the channel model in (4) with a fixed channel matrix. It is possible to obtain lower and upper bounds to the MMSE by relying on the use of a genie based estimator and a (sub-optimal) Euclidean distance based estimator, respectively. The bounds, which we express in terms of the squared pairwise Euclidean distances  $d_{ij}^2(\mathbf{H})$ ,  $i \neq j$ , rather than the squared minimum Euclidean distance  $d_{\min}^2(\mathbf{H}) = \min_{i \neq j} d_{ij}^2(\mathbf{H})$  are, as opposed to the bounds in [16], valid for all signal-to-noise ratios. This aspect is particularly relevant because the bounds to the average value of the MMSE follow from the bounds to the MMSE by averaging over the fading statistics.

*Lemma 1:* The MMSE associated with the estimation of the noiseless output given the noisy output of the channel model in (4), for a fixed channel matrix, can be bounded as follows:

$$\text{mmse}_{\text{LB}}(\text{snr}; \mathbf{H}) \leq \text{mmse}(\text{snr}; \mathbf{H}) \leq \text{mmse}_{\text{UB}}(\text{snr}; \mathbf{H}) \quad (13)$$

where the lower and upper bounds are given by:

$$\text{mmse}_{\text{LB}}(\text{snr}; \mathbf{H}) = \frac{1}{4M(M-1)} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M d_{ij}^2(\mathbf{H}) \cdot \frac{1}{2} \cdot \text{erfc} \left( \sqrt{\frac{d_{ij}^2(\mathbf{H}) \text{snr}}{4}} \right) \quad (14)$$

$$\text{mmse}_{\text{UB}}(\text{snr}; \mathbf{H}) = \frac{1}{M} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M d_{ij}^2(\mathbf{H}) \cdot \frac{1}{2} \cdot \text{erfc} \left( \sqrt{\frac{d_{ij}^2(\mathbf{H}) \text{snr}}{4}} \right) \quad (15)$$

*Proof:* See Appendix A. ■

The lower and upper bounds to the mutual information are obtained by using the upper and lower bounds to the MMSE, respectively, in the integral form of the relationship between the mutual information and the MMSE given

by [14]:<sup>1</sup>

$$I(\text{snr}; \mathbf{H}) = \log M - \int_{\text{snr}}^{\infty} \text{mmse}(\xi; \mathbf{H}) d\xi \quad (16)$$

The bounds, which we also express in terms of the squared pairwise Euclidean distances  $d_{ij}^2(\mathbf{H})$ ,  $i \neq j$ , rather than the squared minimum Euclidean distance  $d_{\min}^2(\mathbf{H}) = \min_{i \neq j} d_{ij}^2(\mathbf{H})$  are, as opposed to the bounds in [16], also valid for all signal-to-noise ratios.

*Lemma 2:* The mutual information between the input and the output of the channel model in (4), for a fixed channel matrix, can be bounded as follows:

$$I_{\text{LB}}(\text{snr}; \mathbf{H}) \leq I(\text{snr}; \mathbf{H}) \leq I_{\text{UB}}(\text{snr}; \mathbf{H}) \quad (17)$$

where the upper and lower bounds are given by:

$$I_{\text{LB}}(\text{snr}; \mathbf{H}) = \log M - \frac{1}{M} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M 2 \cdot e^{-\frac{d_{ij}^2(\mathbf{H}) \text{snr}}{4}} \quad (18)$$

$$I_{\text{UB}}(\text{snr}; \mathbf{H}) = \log M - \frac{1}{M(M-1)} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \frac{1}{4} \cdot \text{erfc} \left( \sqrt{\frac{d_{ij}^2(\mathbf{H}) \text{snr}}{4}} \right) \quad (19)$$

*Proof:* See Appendix B. ■

Let us now consider the channel model in (4) with a random channel matrix. It is possible to obtain lower and upper bounds to the average value of the MMSE by averaging over the fading statistics the lower and upper bounds to the MMSE as follows:

$$\overline{\text{mmse}}_{\text{LB}}(\text{snr}) = \mathbb{E}_{\mathbf{H}} \{ \text{mmse}_{\text{LB}}(\text{snr}; \mathbf{H}) \} = \int \text{mmse}_{\text{LB}}(\text{snr}; \mathbf{H}) \cdot p_{\mathbf{H}}(\mathbf{H}) d\mathbf{H} \quad (20)$$

$$\overline{\text{mmse}}_{\text{UB}}(\text{snr}) = \mathbb{E}_{\mathbf{H}} \{ \text{mmse}_{\text{UB}}(\text{snr}; \mathbf{H}) \} = \int \text{mmse}_{\text{UB}}(\text{snr}; \mathbf{H}) \cdot p_{\mathbf{H}}(\mathbf{H}) d\mathbf{H} \quad (21)$$

where  $p_{\mathbf{H}}(\cdot)$  represents the distribution of  $\mathbf{H}$ . The determination of insightful closed-form expressions for the integrals in (20) and (21) is not simple. Consequently, rather than attempt to solve the integrals, we will exploit results from asymptotic analysis to determine the asymptotic expansion of the integrals. The asymptotic expansions, which lead to considerable insight, are very useful in the regime of high snr.

*Lemma 3:* Assume that  $p_{d_{ij}^2}^{(n)}(d_{ij}^2)$ ,  $n = 0, 1, 2, \dots$  are continuous and integrable in  $[0, \infty)$ . Then, the average value of the MMSE associated with the estimation of the noiseless output given the noisy output of the channel model in (4), for a random channel matrix, can be bounded as follows:

$$\overline{\text{mmse}}_{\text{LB}}(\text{snr}) \leq \overline{\text{mmse}}(\text{snr}) \leq \overline{\text{mmse}}_{\text{UB}}(\text{snr}) \quad (22)$$

<sup>1</sup>This representation assumes that the matrix  $\mathbf{H}$  is non-singular. If the matrix  $\mathbf{H}$  is singular then  $I(\text{snr}; \mathbf{H}) = H(\mathbf{H}\mathbf{x}) - \int_{\text{snr}}^{\infty} \text{mmse}(\xi; \mathbf{H}) d\xi$ .



where the asymptotic expansion as  $\text{snr} \rightarrow \infty$  of the lower and upper bounds are given by:

$$\overline{\text{mmse}}_{\text{LB}}(\text{snr}) = \sum_{n=0}^N \frac{1}{\text{snr}^{n+2}} \cdot k_{\text{LB}_{n+1}} \cdot \left( \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{\mathbf{d}_{ij}^2}^{(n)}(0) \right) + \mathcal{O}\left(\frac{1}{\text{snr}^{N+3}}\right), \quad N = 0, 1, \dots \quad (23)$$

$$\overline{\text{mmse}}_{\text{UB}}(\text{snr}) = \sum_{n=0}^N \frac{1}{\text{snr}^{n+2}} \cdot k_{\text{UB}_{n+1}} \cdot \left( \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{\mathbf{d}_{ij}^2}^{(n)}(0) \right) + \mathcal{O}\left(\frac{1}{\text{snr}^{N+3}}\right), \quad N = 0, 1, \dots \quad (24)$$

and

$$k_{\text{LB}_n} = \frac{1}{2M(M-1)} \cdot \frac{n \cdot 4^n}{\sqrt{\pi}} \cdot \frac{\Gamma(n+3/2)}{\Gamma(n+2)} \quad (25)$$

$$k_{\text{UB}_n} = \frac{2}{M} \cdot \frac{n \cdot 4^n}{\sqrt{\pi}} \cdot \frac{\Gamma(n+3/2)}{\Gamma(n+2)} \quad (26)$$

and  $\Gamma(\cdot)$  is the Gamma function.

*Proof:* See Appendix C. ■

The lower and upper bounds to the average value of the mutual information are obtained by using the upper and lower bounds to the average value of the MMSE, respectively, in the integral form of the relationship between the average mutual information and the average MMSE given by [15]:<sup>2</sup>

$$\bar{\mathcal{I}}(\text{snr}) = \log M - \int_{\text{snr}}^{\infty} \overline{\text{mmse}}(\xi) d\xi \quad (27)$$

Once again, the determination of insightful closed-form expressions for the integrals is not simple. Consequently, rather than attempt to solve the integrals, we will also exploit results from asymptotic analysis to determine the asymptotic expansions of the integrals.

*Lemma 4:* Assume that  $p_{\mathbf{d}_{ij}^2}^{(n)}(\mathbf{d}_{ij}^2)$ ,  $n = 0, 1, 2, \dots$  are continuous and integrable in  $[0, \infty)$ . Then, the average value of the mutual information between the input and the output of the channel model in (4), for a random channel matrix, can be bounded as follows:

$$\bar{\mathcal{I}}_{\text{LB}}(\text{snr}) \leq \bar{\mathcal{I}}(\text{snr}) \leq \bar{\mathcal{I}}_{\text{UB}}(\text{snr}) \quad (28)$$

where the asymptotic expansion as  $\text{snr} \rightarrow \infty$  of the lower and upper bounds are given by:

$$\bar{\mathcal{I}}_{\text{LB}}(\text{snr}) = \log M - \sum_{n=0}^N \frac{1}{\text{snr}^{n+1}} \cdot k'_{\text{LB}_{n+1}} \cdot \left( \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{\mathbf{d}_{ij}^2}^{(n)}(0) \right) + \mathcal{O}\left(\frac{1}{\text{snr}^{N+2}}\right), \quad N = 0, 1, \dots \quad (29)$$

$$\bar{\mathcal{I}}_{\text{UB}}(\text{snr}) = \log M - \sum_{n=0}^N \frac{1}{\text{snr}^{n+1}} \cdot k'_{\text{UB}_{n+1}} \cdot \left( \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{\mathbf{d}_{ij}^2}^{(n)}(0) \right) + \mathcal{O}\left(\frac{1}{\text{snr}^{N+2}}\right), \quad N = 0, 1, \dots \quad (30)$$

<sup>2</sup>This representation assumes that the matrix  $\mathbf{H}$  is non-singular with probability equal to one. If the matrix  $\mathbf{H}$  is not non-singular with probability equal to one then  $\bar{\mathcal{I}}(\text{snr}) = \mathbb{E}_{\mathbf{H}} \{ \mathcal{I}(\mathbf{H}\mathbf{x}|\mathbf{H}) \} - \int_{\text{snr}}^{\infty} \overline{\text{mmse}}(\xi) d\xi$ .

and

$$k'_{\text{LB}_n} = \frac{2}{M} \cdot \frac{4^n}{\sqrt{\pi}} \cdot \frac{\Gamma(n+3/2)}{\Gamma(n+2)} \quad (31)$$

$$k'_{\text{UB}_n} = \frac{1}{2M(M-1)} \cdot \frac{4^n}{\sqrt{\pi}} \cdot \frac{\Gamma(n+3/2)}{\Gamma(n+2)} \quad (32)$$

and  $\Gamma(\cdot)$  is the Gamma function.

*Proof:* See Appendix D. ■

It is important to remark that, in order to determine the asymptotic expansions of the lower and upper bounds to the average value of the MMSE and the average value of the mutual information, we assume that the functions  $p_{d_{ij}^2}^{(n)}(d_{ij}^2)$ ,  $n = 0, 1, 2, \dots$  are continuous and integrable in  $[0, \infty)$ . We verify the assumption for the most common fading channel models, including Rayleigh and Ricean fading models, in subsequent sections.<sup>3</sup>

The asymptotic expansions in Lemmas 3 and 4 are the basis of the characterization of the asymptotic behavior of the average value of the MMSE and the average value of the mutual information in the regime of high snr. Let us define the integer  $d \geq 1$  as follows:

$$d = 1 + \min \left\{ n \in \mathbb{N}_0 : \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(n)}(0) \neq 0 \right\} \quad (33)$$

We now seek to define the rate at which the average value of the MMSE tends to its limit as  $\text{snr} \rightarrow \infty$  as well as the rate at which the average value of the mutual information tends to its limit as  $\text{snr} \rightarrow \infty$ , i.e.,

$$- \lim_{\text{snr} \rightarrow \infty} \frac{\log \overline{\text{mmse}}(\text{snr})}{\log \text{snr}} \quad (34)$$

and

$$- \lim_{\text{snr} \rightarrow \infty} \frac{\log(\log M - \bar{I}(\text{snr}))}{\log \text{snr}} \quad (35)$$

These rates are trivially bounded as follows:

$$- \lim_{\text{snr} \rightarrow \infty} \frac{\log \overline{\text{mmse}}_{\text{UB}}(\text{snr})}{\log \text{snr}} \leq - \lim_{\text{snr} \rightarrow \infty} \frac{\log \overline{\text{mmse}}(\text{snr})}{\log \text{snr}} \leq - \lim_{\text{snr} \rightarrow \infty} \frac{\log \overline{\text{mmse}}_{\text{LB}}(\text{snr})}{\log \text{snr}} \quad (36)$$

and

$$- \lim_{\text{snr} \rightarrow \infty} \frac{\log(\log M - \bar{I}_{\text{LB}}(\text{snr}))}{\log \text{snr}} \leq - \lim_{\text{snr} \rightarrow \infty} \frac{\log(\log M - \bar{I}(\text{snr}))}{\log \text{snr}} \leq - \lim_{\text{snr} \rightarrow \infty} \frac{\log(\log M - \bar{I}_{\text{UB}}(\text{snr}))}{\log \text{snr}} \quad (37)$$

By capitalizing on the asymptotic expansions embodied in Lemmas 3 and 4 it is possible to write as  $\text{snr} \rightarrow \infty$ :

$$\log \overline{\text{mmse}}_{\text{LB}}(\text{snr}) = \log \left( k_{\text{LB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) \right) - (d+1) \log \text{snr} + \mathcal{O}\left(\frac{1}{\text{snr}}\right) \quad (38)$$

<sup>3</sup>We will see that the fact that  $p_{d_{ij}^2}^{(n)}(d_{ij}^2)$ ,  $n = 1, 2, \dots$  may be discontinuous at zero is immaterial.

$$\log \overline{\text{mmse}}_{\text{UB}}(\text{snr}) = \log \left( k_{\text{UB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) \right) - (d+1) \log \text{snr} + \mathcal{O} \left( \frac{1}{\text{snr}} \right) \quad (39)$$

and

$$\log (\log M - \bar{\text{I}}_{\text{LB}}(\text{snr})) = \log \left( k'_{\text{LB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) \right) - d \log \text{snr} + \mathcal{O} \left( \frac{1}{\text{snr}} \right) \quad (40)$$

$$\log (\log M - \bar{\text{I}}_{\text{UB}}(\text{snr})) = \log \left( k'_{\text{UB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) \right) - d \log \text{snr} + \mathcal{O} \left( \frac{1}{\text{snr}} \right) \quad (41)$$

so that

$$-\lim_{\text{snr} \rightarrow \infty} \frac{\log \overline{\text{mmse}}(\text{snr})}{\log \text{snr}} = -\lim_{\text{snr} \rightarrow \infty} \frac{\log \overline{\text{mmse}}_{\text{LB}}(\text{snr})}{\log \text{snr}} = -\lim_{\text{snr} \rightarrow \infty} \frac{\log \overline{\text{mmse}}_{\text{UB}}(\text{snr})}{\log \text{snr}} = d+1 \quad (42)$$

and

$$-\lim_{\text{snr} \rightarrow \infty} \frac{\log (\log M - \bar{\text{I}}(\text{snr}))}{\log \text{snr}} = -\lim_{\text{snr} \rightarrow \infty} \frac{\log (\log M - \bar{\text{I}}_{\text{LB}}(\text{snr}))}{\log \text{snr}} = -\lim_{\text{snr} \rightarrow \infty} \frac{\log (\log M - \bar{\text{I}}_{\text{UB}}(\text{snr}))}{\log \text{snr}} = d \quad (43)$$

We also seek to define, in addition to the rates at which the average value of the MMSE and the average value of the mutual information tend to their infinite-snr values, a finer characterization of the high-snr asymptotic behavior.

Towards this end, we define the quantities:

$$\bar{\epsilon}_d = \limsup_{\text{snr} \rightarrow \infty} \text{snr}^{d+1} \cdot \overline{\text{mmse}}(\text{snr}) \quad (44)$$

$$\underline{\epsilon}_d = \liminf_{\text{snr} \rightarrow \infty} \text{snr}^{d+1} \cdot \overline{\text{mmse}}(\text{snr}) \quad (45)$$

and

$$\bar{\epsilon}'_d = \limsup_{\text{snr} \rightarrow \infty} \text{snr}^d \cdot (\log M - \bar{\text{I}}(\text{snr})) \quad (46)$$

$$\underline{\epsilon}'_d = \liminf_{\text{snr} \rightarrow \infty} \text{snr}^d \cdot (\log M - \bar{\text{I}}(\text{snr})) \quad (47)$$

Note that these quantities, in view of the asymptotic expansions embodied in Lemmas 3 and 4, can be bounded as follows:

$$k_{\text{LB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) \leq \underline{\epsilon}_d \leq \bar{\epsilon}_d \leq k_{\text{UB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) \quad (48)$$

and

$$k'_{\text{UB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) \leq \underline{\epsilon}'_d \leq \bar{\epsilon}'_d \leq k'_{\text{LB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) \quad (49)$$

Note also that the bounds in (48) and (49) are finite for key fading models (see Sections IV and V).

The following Theorems, which are based on these considerations, characterize the asymptotic behavior of the average value of the MMSE and the average value of the mutual information in the regime of high snr.

*Theorem 1:* Assume that  $p_{d_{ij}^2}^{(n)}(d_{ij}^2)$ ,  $n = 0, 1, 2, \dots$  are continuous and integrable in  $[0, \infty)$ . Then, in the regime of high-snr the average value of the MMSE associated with the estimation of the noiseless output given the noisy output of the channel model in (4) can be expanded as follows:

$$\overline{\text{mmse}}(\text{snr}) = \epsilon_d(\text{snr}) \cdot \frac{1}{\text{snr}^{d+1}} + \mathcal{O}\left(\frac{1}{\text{snr}^{d+2}}\right) \quad (50)$$

where  $\epsilon_d(\text{snr})$  is a piecewise infinitely differentiable function such that:

$$k_{\text{LB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) \leq \epsilon_d(\text{snr}) \leq k_{\text{UB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) \quad (51)$$

in the interval  $[\text{snr}_0, \infty)$  for a sufficiently high value of  $\text{snr}_0$  and

$$\limsup_{\text{snr} \rightarrow \infty} \epsilon_d(\text{snr}) = \bar{\epsilon}_d \quad \text{and} \quad \liminf_{\text{snr} \rightarrow \infty} \epsilon_d(\text{snr}) = \underline{\epsilon}_d \quad (52)$$

*Proof:* See Appendix E. ■

*Theorem 2:* Assume that  $p_{d_{ij}^2}^{(n)}(d_{ij}^2)$ ,  $n = 0, 1, 2, \dots$  are continuous and integrable in  $[0, \infty)$ . Then, in the regime of high-snr the average value of the mutual information between the input and the output of the channel model in (4) can be expanded as follows:

$$\bar{I}(\text{snr}) = \log M - \epsilon'_d(\text{snr}) \cdot \frac{1}{\text{snr}^d} + \mathcal{O}\left(\frac{1}{\text{snr}^{d+1}}\right) \quad (53)$$

where  $\epsilon'_d(\text{snr})$  is a piecewise infinitely differentiable function such that:

$$k'_{\text{UB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) \leq \epsilon'_d(\text{snr}) \leq k'_{\text{LB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) \quad (54)$$

in the interval  $[\text{snr}_0, \infty)$  for a sufficiently high value of  $\text{snr}_0$  and

$$\limsup_{\text{snr} \rightarrow \infty} \epsilon'_d(\text{snr}) = \bar{\epsilon}'_d \quad \text{and} \quad \liminf_{\text{snr} \rightarrow \infty} \epsilon'_d(\text{snr}) = \underline{\epsilon}'_d \quad (55)$$

*Proof:* See also Appendix E. ■

Note that if

$$\bar{\epsilon}_d = \underline{\epsilon}_d = \lim_{\text{snr} \rightarrow \infty} \text{snr}^{d+1} \cdot \overline{\text{mmse}}(\text{snr}) \triangleq \epsilon_d \quad (56)$$

and

$$\bar{\epsilon}'_d = \underline{\epsilon}'_d = \lim_{\text{snr} \rightarrow \infty} \text{snr}^d \cdot (\log M - \bar{I}(\text{snr})) \triangleq \epsilon'_d \quad (57)$$

then the asymptotic expansions exposed in Theorems 1 and 2 reduce immediately to:

$$\overline{\text{mmse}}(\text{snr}) = \epsilon_d \cdot \frac{1}{\text{snr}^{d+1}} + \mathcal{O}\left(\frac{1}{\text{snr}^{d+2}}\right) \quad (58)$$

and

$$\bar{I}(\text{snr}) = \log M - \epsilon'_d \cdot \frac{1}{\text{snr}^d} + \mathcal{O}\left(\frac{1}{\text{snr}^{d+1}}\right) \quad (59)$$

Therefore, in view of (27), (58) and (59) it is possible to establish the relation:

$$\epsilon'_d = \lim_{\text{snr} \rightarrow \infty} \text{snr}^d \cdot (\log M - \bar{I}(\text{snr})) = \frac{1}{d} \cdot \lim_{\text{snr} \rightarrow \infty} \text{snr}^{d+1} \cdot \overline{\text{mmse}}(\text{snr}) = \frac{1}{d} \cdot \epsilon_d \quad (60)$$

so that, as expected, the scaling constants that define the high-snr asymptotics of the average value of the MMSE and the average value of the mutual information are also related. Note also that the quantities  $\epsilon_d$  and  $\epsilon'_d$  represent a generalization of the MMSE dimension, which, when it exists, defines the high-snr asymptotics of the MMSE of a random variable observed in zero-mean unit-variance Gaussian noise [30]. Numerical results suggest that the limits in (56) and (57) exist for common multiple-antenna fading coherent channel models driven by arbitrary equiprobable discrete inputs, thereby justifying the use of the expansions in (58) and (59) to characterize the asymptotic behavior.

It is also important to note that in general we can only bound  $\epsilon_d$  and  $\epsilon'_d$  as in (48) and (49), respectively, rather than compute their exact values. The upper and lower bounds to  $\epsilon_d$  and  $\epsilon'_d$  differ by a factor of  $4 \cdot (M - 1)$ , and so become increasingly loose for multi-dimensional constellations with high cardinality. The lower bound to  $\epsilon_d$  and the upper bound to  $\epsilon'_d$ , which are due to the genie based estimator, are considerably loose.<sup>4</sup> In contrast, the upper bound to  $\epsilon_d$  and the lower bound to  $\epsilon'_d$ , which are due to the Euclidean distance based estimator, are considerably tighter. In general, the use of the bounds to  $\epsilon_d$  and  $\epsilon'_d$  in the expansions in (58) and (59), respectively, leads to bounds to the true asymptotic expansions. The bounds to the asymptotic expansions and the true asymptotic expansions differ only by an offset in signal-to-noise ratio. The offset in dB between the high-snr expansion of the average value of the MMSE in (58) and the expansions that use the lower and upper bounds to the value of  $\epsilon_d$  is given respectively by:

$$\Delta_{\text{LB}}|_{\text{dB}} = 10 \cdot \log_{10} \left( \frac{k_{\text{LB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0)}{\epsilon_d} \right)^{\frac{1}{d+1}} \quad (61)$$

and

$$\Delta_{\text{UB}}|_{\text{dB}} = 10 \cdot \log_{10} \left( \frac{k_{\text{UB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0)}{\epsilon_d} \right)^{\frac{1}{d+1}} \quad (62)$$

<sup>4</sup>It is possible to construct genie based estimators that lead to tighter bounds. The current genie supplies the receiver with a pair of input vectors for each transmit vector, the true input vector and any of the other input vectors with equal probability. A more appropriate genie supplies the receiver with the true input vector and another suitable input vector for each transmit vector. In particular, in the canonical i.i.d. Rayleigh fading coherent channel, Theorem 3 suggests that the genie ought to minimize as much as possible the sum of the inverse of the squared Euclidean distances between the pairs of supplied input vectors. It is also important to guarantee, in addition, that such a genie construction leads indeed to a lower bound to the MMSE and hence an upper bound to the mutual information. This is met, for example, by constructing a genie that also ensures that the receiver sees all the input vectors with equal probability. We do not pursue this issue further, because the Euclidean distance based estimator upper bound to  $\epsilon_d$  and lower bound to  $\epsilon'_d$  are still tighter than the new genie based estimator lower bound to  $\epsilon_d$  and upper bound to  $\epsilon'_d$ .

TABLE I  
SIGNAL-TO-NOISE RATIO OFFSET BETWEEN THE TRUE ASYMPTOTIC EXPANSIONS AND THE BOUNDS TO THE ASYMPTOTIC EXPANSIONS  
FOR MULTIPLE-ANTENNA RAYLEIGH FADING CHANNELS DRIVEN BY 16-QAM INPUTS

$(n_t, n_r)$	$\Delta_{\text{LB}} _{\text{dB}}$	$\Delta_{\text{UB}} _{\text{dB}}$	$\Delta'_{\text{LB}} _{\text{dB}}$	$\Delta'_{\text{UB}} _{\text{dB}}$
(1,1)	-6.9 dB	2.0 dB	3.9 dB	-13.8 dB
(1,2)	-5.0 dB	0.9 dB	1.4 dB	-7.5 dB
(1,3)	-4.0 dB	0.5 dB	0.7 dB	-5.3 dB
(2,2)	-9.2 dB	0.8 dB	1.2 dB	-13.8 dB
(3,3)	-9.8 dB	0.8 dB	1.0 dB	-13.0 dB

TABLE II  
SIGNAL-TO-NOISE RATIO OFFSET BETWEEN THE TRUE ASYMPTOTIC EXPANSIONS AND THE BOUNDS TO THE ASYMPTOTIC EXPANSIONS  
FOR SINGLE-ANTENNA RAYLEIGH FADING CHANNELS DRIVEN BY VARIOUS INPUTS

	$\Delta_{\text{LB}} _{\text{dB}}$	$\Delta_{\text{UB}} _{\text{dB}}$	$\Delta'_{\text{LB}} _{\text{dB}}$	$\Delta'_{\text{UB}} _{\text{dB}}$
BPSK	-2.4 dB	0.6 dB	1.1 dB	-4.9 dB
QPSK	-4.3 dB	1.1 dB	2.2 dB	-8.6 dB
16-QAM	-6.9 dB	2.0 dB	3.9 dB	-13.9 dB
64-QAM	-9.4 dB	2.6 dB	5.3 dB	-18.7 dB
256-QAM	-11.7 dB	3.3 dB	6.6 dB	-23.5 dB

In turn, the offset in dB between the high-snr expansion of the average value of the mutual information in (59) and the expansions that use the lower and upper bounds to the value of  $\epsilon'_d$  is given respectively by:

$$\Delta'_{\text{LB}}|_{\text{dB}} = 10 \cdot \log_{10} \left( \frac{k'_{\text{LB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0)}{\epsilon'_d} \right)^{\frac{1}{d}} \quad (63)$$

and

$$\Delta'_{\text{UB}}|_{\text{dB}} = 10 \cdot \log_{10} \left( \frac{k'_{\text{UB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0)}{\epsilon'_d} \right)^{\frac{1}{d}} \quad (64)$$

Tables I and II show the signal-to-noise ratio offset values for particular system configurations. We observe that indeed  $\Delta_{\text{UB}}|_{\text{dB}}$  or  $\Delta'_{\text{LB}}|_{\text{dB}}$  are much lower than  $\Delta_{\text{LB}}|_{\text{dB}}$  or  $\Delta'_{\text{UB}}|_{\text{dB}}$ , respectively, which is a manifestation of the fact that the upper bound to  $\epsilon_d$  and the lower bound to  $\epsilon'_d$  are considerably tighter than the lower bound to  $\epsilon_d$  and the upper bound to  $\epsilon'_d$ . We also observe that the signal-to-noise ratio offset values  $\Delta_{\text{UB}}|_{\text{dB}}$  and  $\Delta'_{\text{LB}}|_{\text{dB}}$  decrease with the increase in the number of antennas and, as expected, increase with the increase in the multi-dimensional constellation cardinality. As an example, it is interesting to note that  $\Delta_{\text{UB}}|_{\text{dB}} = 0.8$  dB and  $\Delta'_{\text{LB}}|_{\text{dB}} = 1.2$  dB for a two-transmit two-receive antenna canonical i.i.d. Rayleigh fading coherent channel driven by 16-QAM inputs.

Figures 1, 2 and 3 depict the average value of the MMSE and the average value of the mutual information and

the respective asymptotic expansions for key multiple-antenna systems and models.<sup>5</sup> We observe a very reasonable match between the exact curve, obtained through Monte Carlo simulations, and the asymptotic expansion in the high-snr regime. We also observe that the asymptotic expansions characterize perfectly the rates at which the average MMSE and the average mutual information approach their infinite snr values. In general, the use of the high-snr asymptotic expansions of the average MMSE and the average mutual information results in errors that relate to the fact that: *i*) one can only compute the leading term rather than other possible higher-order terms in the high-snr asymptotic expansions; and *ii*) one can only compute analytically bounds to  $\epsilon_d$  and  $\epsilon'_d$ . The first consideration leads to the differences between the asymptotic expansions and the exact curve in Figures 1, 2 and 3. The second consideration leads to the additional signal-to-noise ratio offsets in Tables I and II. Nonetheless, it will be shown that the asymptotic expansions still produce very accurate designs in Section VI.

The regime of validity of the asymptotic expansions in certain fading models also deserves more attention. The Ricean fading model is an intriguing case study. For example, it is well known that with the increase in the  $K$ -factor a Ricean fading channel will approach an AWGN-like channel, so that one would expect the average MMSE and the average mutual information to tend to their infinite-snr values at a much faster rate than  $d + 1$  and  $d$ , respectively, as put forth in Theorems 1 and 2. Figures 4 and 5 overcome this apparent contradiction by showing that the average MMSE and the average mutual information do indeed tend to their infinite-snr values at rates  $d + 1$  and  $d$ , respectively. However, the influence of such terms in the expansions only shows up at extremely high signal-to-noise ratios; other possible terms in the expansion – which mimic more closely the AWGN-like behavior – play a more prominent role at moderate and high signal-to-noise ratios. Obviously, in such circumstances, typically associated with larger  $K$ -factors, the expansions will be of little practical use because they fail to characterize the behavior of the quantities in regimes of interest.

Another interesting issue relates to possible generalizations of the expansions in (58) and (59). In view of the expansions of the upper and lower bounds to the average value of the MMSE as well as the upper and lower bounds to the average value of the mutual information put forth in Lemmas 3 and 4, and the arguments that lead to Theorems 1 and 2, it is also tempting to conjecture that the average MMSE behaves as:

$$\overline{\text{mmse}}(\text{snr}) = \sum_{k=d}^N \epsilon_k \cdot \frac{1}{\text{snr}^{k+1}} + o\left(\frac{1}{\text{snr}^{N+1}}\right) \quad (65)$$

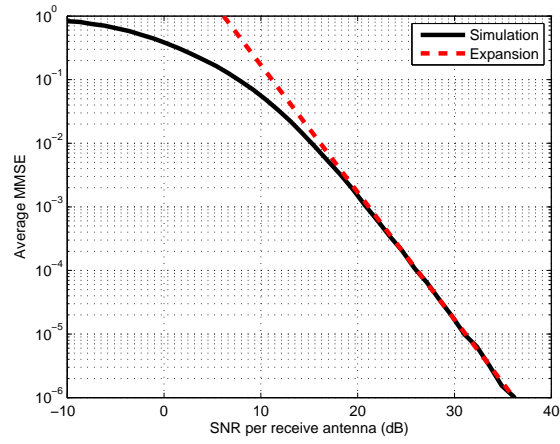
for arbitrary  $N$ , with

$$\epsilon_k = \lim_{\text{snr} \rightarrow \infty} \text{snr}^{k+1} \cdot \left( \overline{\text{mmse}}(\text{snr}) - \sum_{k'=d}^{k-1} \epsilon_{k'} \cdot \frac{1}{\text{snr}^{k'+1}} \right) \quad (66)$$

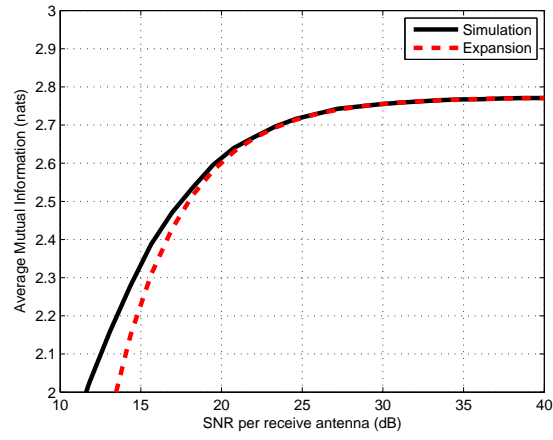
whereas the average mutual information behaves as:

$$\bar{I}(\text{snr}) = \log M - \sum_{k=d}^N \epsilon'_k \cdot \frac{1}{\text{snr}^k} + o\left(\frac{1}{\text{snr}^N}\right) \quad (67)$$

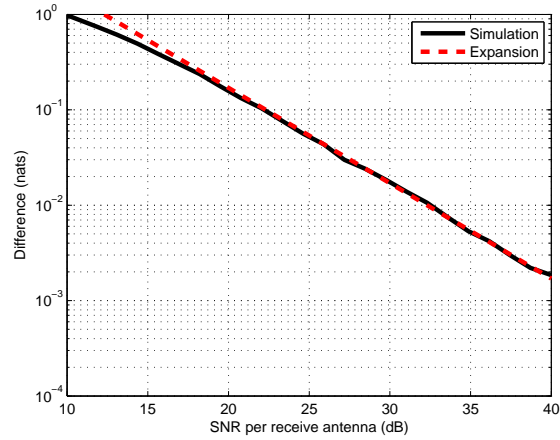
<sup>5</sup>We use the exact values of  $\epsilon_d$  and  $\epsilon'_d$ , obtained through Monte Carlo simulations, in the asymptotic expansions.



(a)



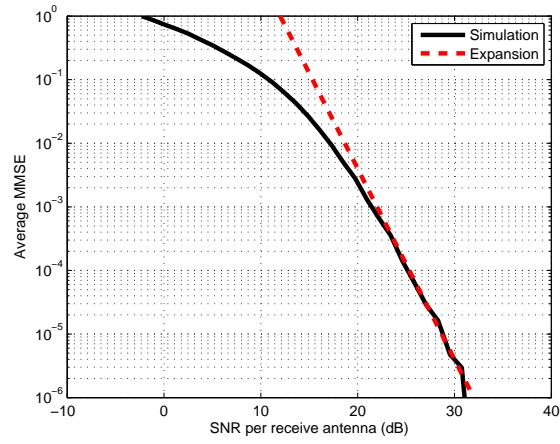
(b)



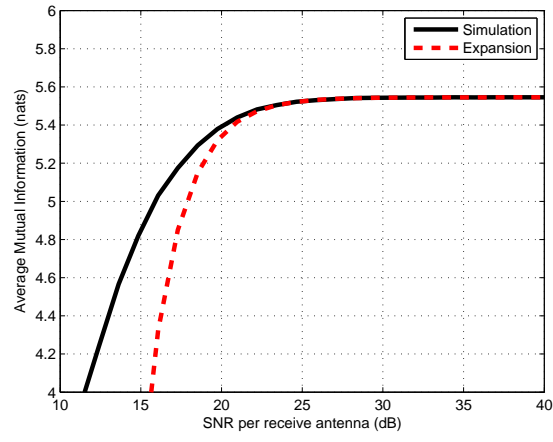
(c)

Fig. 1.  $1 \times 1$  Rayleigh fading coherent channel with a 16-QAM input: a) average MMSE; b) average mutual information; (c) difference between maximum average mutual information and average mutual information.

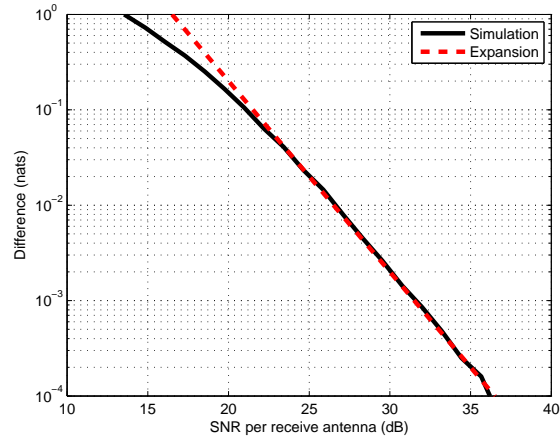




(a)

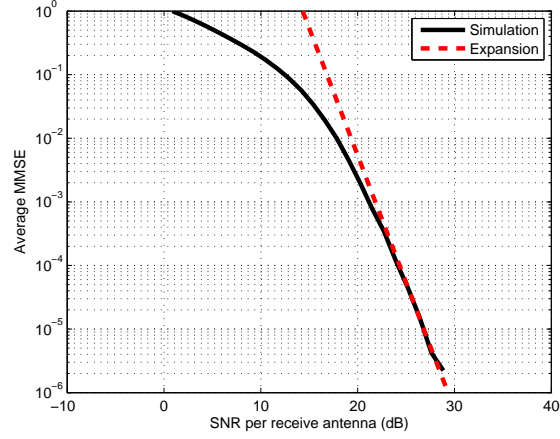


(b)

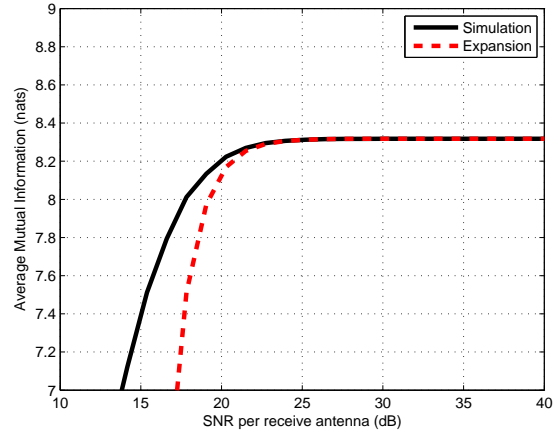


(c)

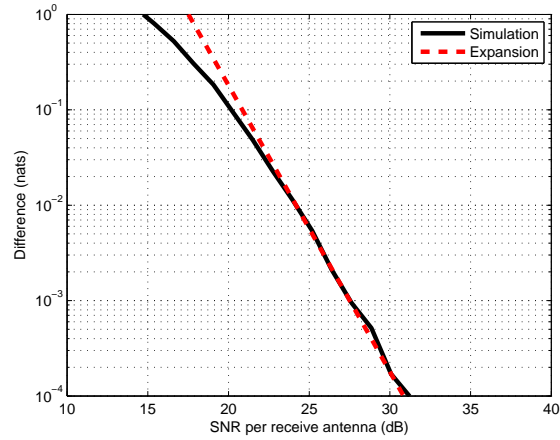
Fig. 2.  $2 \times 2$  Rayleigh fading coherent channel with 16-QAM inputs: a) average MMSE; b) average mutual information; (c) difference between maximum average mutual information and average mutual information.



(a)

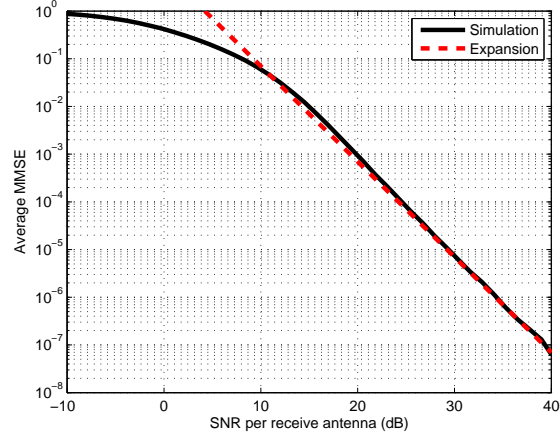


(b)

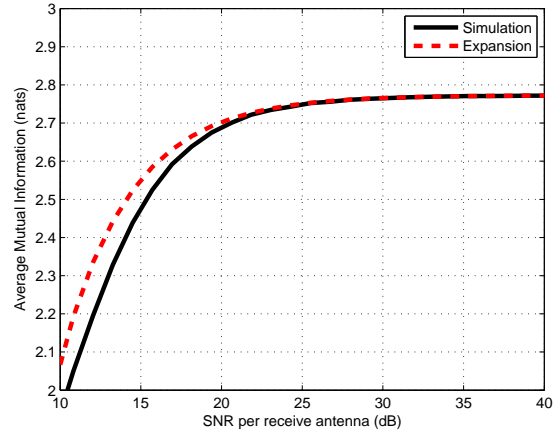


(c)

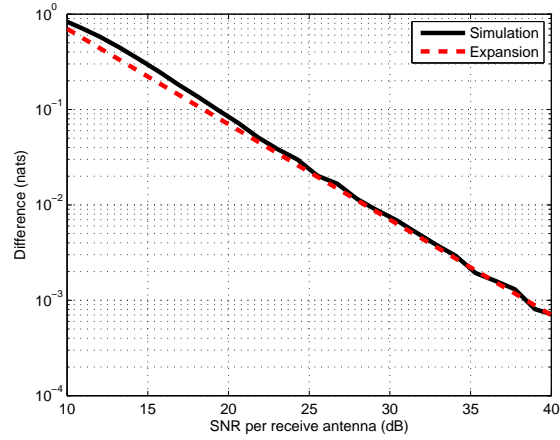
Fig. 3.  $3 \times 3$  Rayleigh fading coherent channel with 16-QAM inputs: a) average MMSE; b) average mutual information; (c) difference between maximum average mutual information and average mutual information.



(a)

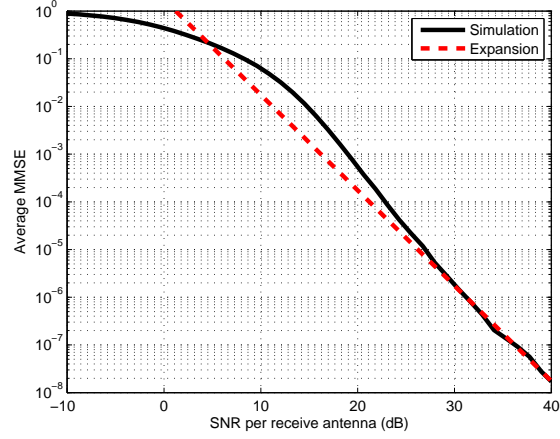


(b)

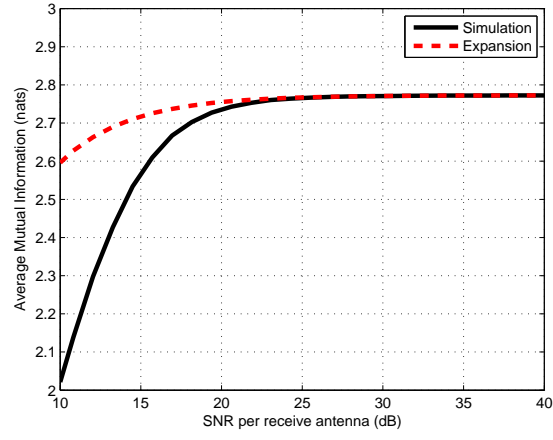


(c)

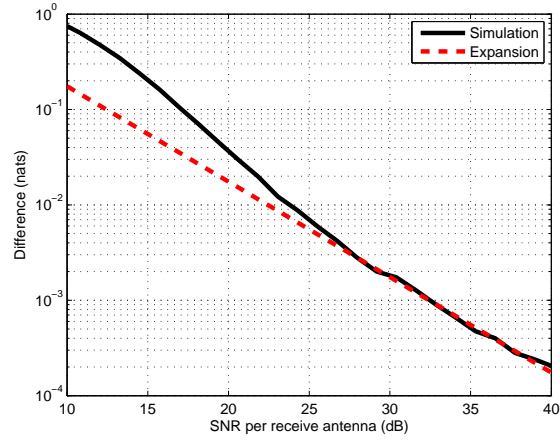
Fig. 4.  $1 \times 1$  Ricean fading coherent channel with a 16-QAM input ( $K = 2$ ,  $\mathbf{H}_0 = \mathbf{I}$ ): a) average MMSE; b) average mutual information; (c) difference between maximum average mutual information and average mutual information.



(a)



(b)



(c)

Fig. 5.  $1 \times 1$  Ricean fading coherent channel with a 16-QAM input ( $K = 4$ ,  $\mathbf{H}_0 = \mathbf{I}$ ): a) average MMSE; b) average mutual information; (c) difference between maximum average mutual information and average mutual information.

for arbitrary  $N$  with

$$\epsilon'_k = \lim_{\text{snr} \rightarrow \infty} \text{snr}^k \cdot \left( \log M - \bar{I}(\text{snr}) - \sum_{k'=d}^{k-1} \epsilon'_{k'} \cdot \frac{1}{\text{snr}^{k'}} \right) \quad (68)$$

Expansions (65) and (67) would then provide a dissection of the high-snr behavior of the average value of the MMSE and the average value of the mutual information of arbitrary equiprobable discrete inputs observed through a multiple-antenna fading coherent channel in Gaussian noise in terms of a series of quantities, namely,  $\epsilon_k$  and  $\epsilon'_k$  for  $k = d, d+1, d+2, \dots$ . Interestingly, though our analytic techniques do not lead to (65) and (67), Figures 6 and 7 show that such expansions provide a more accurate representation of the average value of the MMSE and the average value of the mutual information than the expansions in (58) and (59).

Finally, it is also of interest to characterize the asymptotic behavior of the probability of error in a multiple-antenna fading coherent channel with arbitrary equiprobable discrete inputs in the regime of high snr. We consider upper and lower bounds to the (uncoded) probability of error associated with a maximum likelihood detector for the channel model in (4), for a fixed channel matrix, given by:

$$P_e(\text{snr}; \mathbf{H}) = \Pr(\hat{\mathbf{x}}_{\text{ML}}(\sqrt{\text{snr}} \cdot \mathbf{H}\mathbf{x} + \mathbf{n}) \neq \mathbf{x} | \mathbf{H}) \quad (69)$$

as well as upper and lower bounds to the average value of the (uncoded) probability of error associated with the maximum-likelihood detector for the channel model in (4), for a random channel matrix, given by:

$$\bar{P}_e(\text{snr}) = \mathbb{E}_{\mathbf{H}} \{P_e(\text{snr}; \mathbf{H})\} = \mathbb{E}_{\mathbf{H}} \left\{ \Pr(\hat{\mathbf{x}}_{\text{ML}}(\sqrt{\text{snr}} \cdot \mathbf{H}\mathbf{x} + \mathbf{n}) \neq \mathbf{x} | \mathbf{H}) \right\} \quad (70)$$

where  $\hat{\mathbf{x}}_{\text{ML}}(\sqrt{\text{snr}} \cdot \mathbf{H}\mathbf{x} + \mathbf{n})$  corresponds to the maximum-likelihood detector estimate of the input given the output of the channel model in (4). We also consider the asymptotic expansions of the upper and lower bounds to the average value of the probability of error. Lemmas 5 and 6 summarize the results.

*Lemma 5:* The (uncoded) probability of error associated with maximum likelihood detection in the channel model in (4), for a fixed channel matrix, can be bounded as follows:

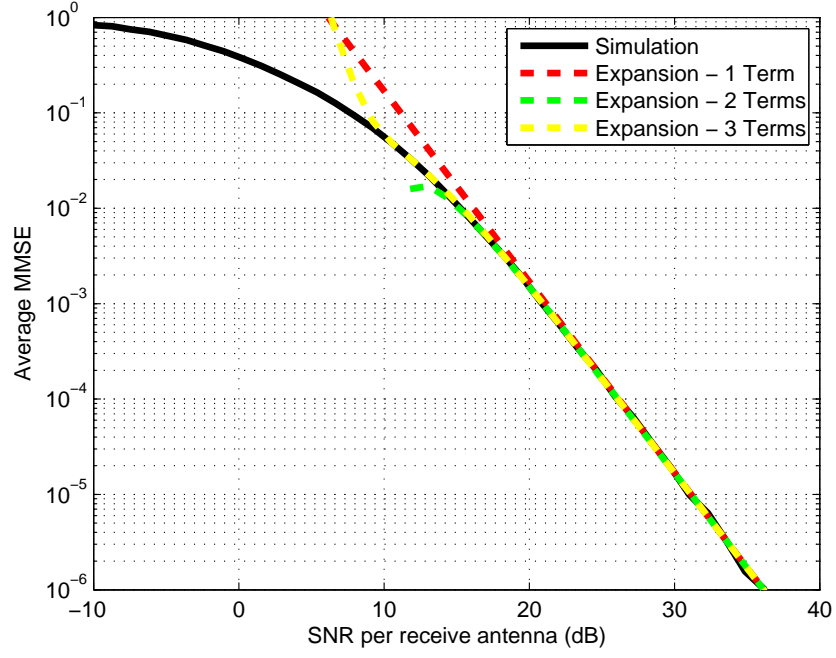
$$P_{\text{eLB}}(\text{snr}; \mathbf{H}) \leq P_e(\text{snr}; \mathbf{H}) \leq P_{\text{eUB}}(\text{snr}; \mathbf{H}) \quad (71)$$

where the lower and upper bounds are given by:

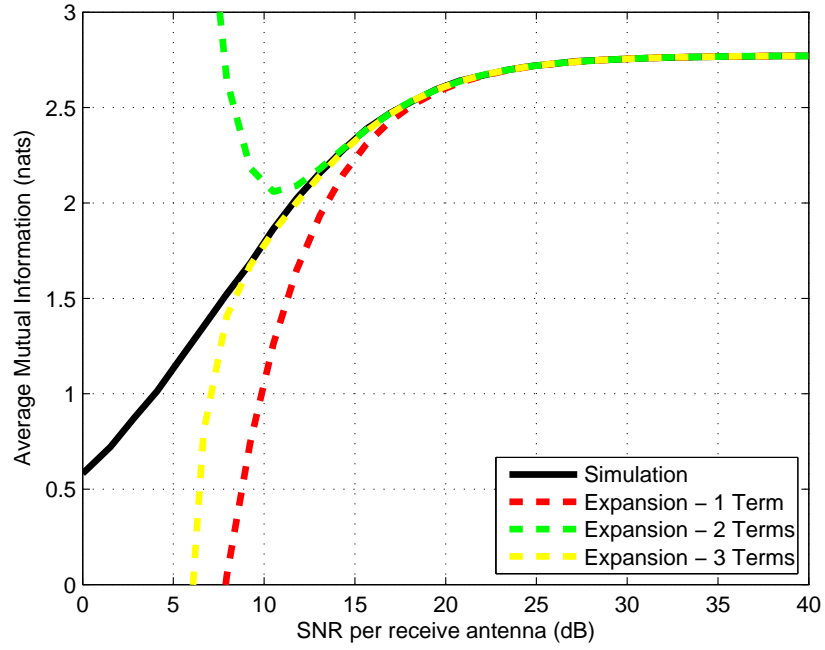
$$P_{\text{eLB}}(\text{snr}; \mathbf{H}) = \frac{1}{M(M-1)} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \frac{1}{2} \cdot \text{erfc} \left( \sqrt{\frac{d_{ij}^2(\mathbf{H}) \text{snr}}{4}} \right) \quad (72)$$

$$P_{\text{eUB}}(\text{snr}; \mathbf{H}) = \frac{1}{M} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \frac{1}{2} \cdot \text{erfc} \left( \sqrt{\frac{d_{ij}^2(\mathbf{H}) \text{snr}}{4}} \right) \quad (73)$$

*Proof:* See Appendix F. ■

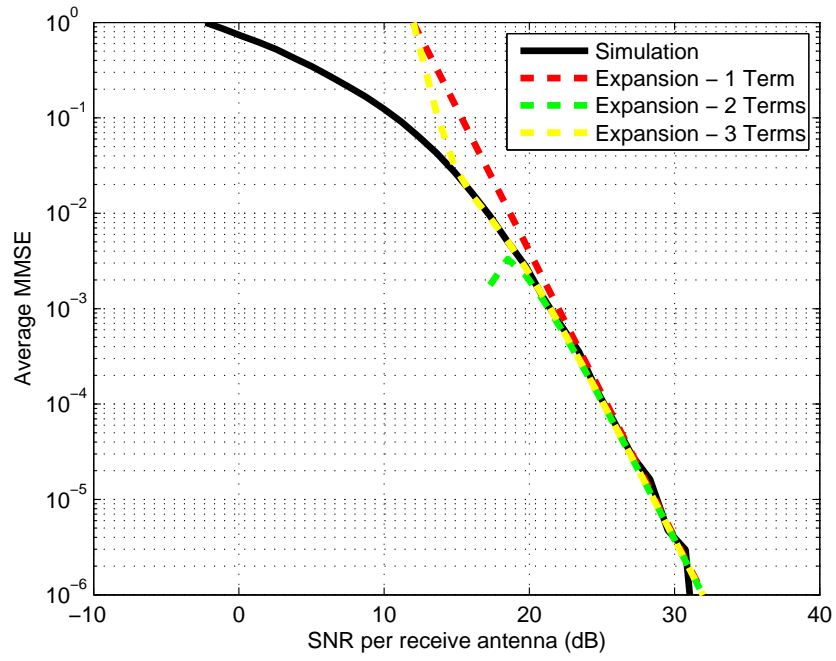


(a)

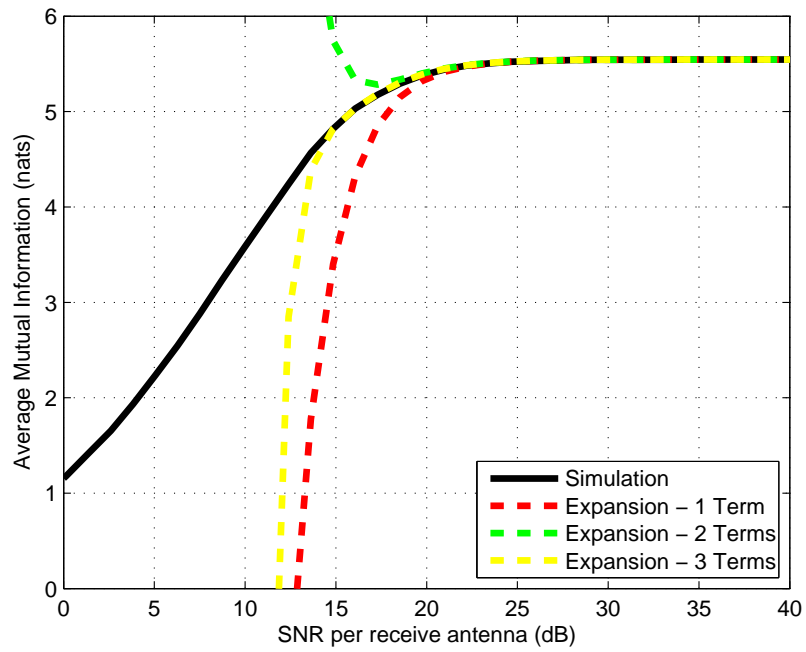


(b)

Fig. 6.  $1 \times 1$  Rayleigh fading coherent channel with a 16-QAM input : a) higher-order expansions of average MMSE; b) higher-order expansions of average mutual information.



(a)



(b)

Fig. 7.  $2 \times 2$  Rayleigh fading coherent channel with 16-QAM inputs : a) higher-order expansions of average MMSE; b) higher-order expansions of average mutual information.

*Lemma 6:* Assume that  $p_{d_{ij}^2}^{(n)}(d_{ij}^2), n = 0, 1, 2, \dots$  are continuous and integrable in  $[0, \infty)$ . Then, the average value of the (uncoded) probability of error associated with maximum likelihood detection in the channel model in (4), for a random channel matrix, can be bounded as follows:

$$\bar{P}_{e_{LB}}(\text{snr}) \leq \bar{P}_e(\text{snr}) \leq \bar{P}_{e_{UB}}(\text{snr}) \quad (74)$$

where the asymptotic expansion as  $\text{snr} \rightarrow \infty$  of the lower and upper bounds are given by:

$$\bar{P}_{e_{LB}}(\text{snr}) = \sum_{n=0}^N \frac{1}{\text{snr}^{n+1}} \cdot k_{LB_{n+1}}'' \cdot \left( \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(n)}(0) \right) + \mathcal{O}\left(\frac{1}{\text{snr}^{N+2}}\right), \quad N = 0, 1, \dots \quad (75)$$

$$\bar{P}_{e_{UB}}(\text{snr}) = \sum_{n=0}^N \frac{1}{\text{snr}^{n+1}} \cdot k_{UB_{n+1}}'' \cdot \left( \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(n)}(0) \right) + \mathcal{O}\left(\frac{1}{\text{snr}^{N+2}}\right), \quad N = 0, 1, \dots \quad (76)$$

and

$$k_{LB_n}'' = \frac{1}{2M(M-1)} \cdot \frac{4^n}{\sqrt{\pi}} \cdot \frac{\Gamma(n+1/2)}{\Gamma(n+1)} \quad (77)$$

$$k_{UB_n}'' = \frac{1}{2M} \cdot \frac{4^n}{\sqrt{\pi}} \cdot \frac{\Gamma(n+1/2)}{\Gamma(n+1)} \quad (78)$$

and  $\Gamma(\cdot)$  is the Gamma function.

*Proof:* See Appendix G. ■

Note that as  $\text{snr} \rightarrow \infty$  the upper and lower bounds to the average value of the probability of error are arbitrarily tight so that:

$$\bar{P}_e(\text{snr}) = \mathcal{O}\left(\frac{1}{\text{snr}^d}\right) \neq o\left(\frac{1}{\text{snr}^d}\right) \quad (79)$$

We immediately recognize the quantity  $d$ , which also appears in the expansions of the average value of the MMSE and the average value of the mutual information in Theorems 1 and 2, respectively, to be the familiar diversity gain given by [31]:

$$d = - \lim_{\text{snr} \rightarrow \infty} \frac{\log \bar{P}_e(\text{snr})}{\log \text{snr}} \quad (80)$$

Interestingly, the high-snr asymptotic behavior of the average value of the minimum mean-squared error, the average value of the mutual information and also the average value of the probability of error exhibit some common features, with the value of the parameter  $d$  and the value of the bounds to  $\epsilon_d(\text{snr})$  and  $\epsilon_d'(\text{snr})$  dictated by the value of the probability density function of the squared pairwise Euclidean distances at zero,  $p_{d_{ij}^2}(0)$ , or the value of their higher-order derivatives at zero,  $p_{d_{ij}^2}^{(n)}(0), n > 0$ . This dependency, which has also been very briefly noticed in [32], is intuitive because one could argue that in fading channels at high-snr the quantities are affected primarily by the value of the probability of arbitrarily close (noiseless) receive vectors due to unfavorable fading realizations, in the same way that in non-fading channels at high-snr the minimum mean-squared error, the mutual information,



and the probability of error are primarily affected by the value of the Euclidean distance between the (noiseless) receive vectors [16]. As a side remark, we observe that the values of  $p_{d_{ij}^2}(0)$  or  $p_{d_{ij}^2}^{(n)}(0), n > 0$ , which depend on the channel statistics as well as the system elements, can then be the basis of the characterization or the optimization of the constrained capacity of multiple-antenna fading coherent channels driven by arbitrary equiprobable discrete inputs in the regime of high snr. Sections IV, V and VI concentrate on such issues.

#### IV. THE CANONICAL I.I.D. RAYLEIGH FADING COHERENT CHANNEL

It is now relevant to characterize the constrained capacity of the canonical i.i.d. Rayleigh fading coherent channel with arbitrary equiprobable discrete inputs in the regime of high-snr. The objective is to understand the effect on the constrained capacity of various system parameters, such as the number of transmit antennas, the number of receive antennas as well as the characteristics of the signalling scheme.

The following Theorem unveils the high-snr behavior of the constrained capacity. We represent the squared pairwise Euclidean distance between two arbitrary transmit vectors  $\mathbf{x}_i$  and  $\mathbf{x}_j$  by  $\bar{d}_{ij}^2 = \|\mathbf{x}_i - \mathbf{x}_j\|^2$ .

*Theorem 3:* In the regime of high-snr, the constrained capacity of the canonical i.i.d. multiple-transmit-multiple-receive antenna Rayleigh fading coherent channel with arbitrary equiprobable discrete inputs obeys:

$$\bar{I}(\text{snr}) = \log M - \epsilon'_{n_r}(\text{snr}) \cdot \frac{1}{\text{snr}^{n_r}} + \mathcal{O}\left(\frac{1}{\text{snr}^{n_r+1}}\right) \quad (81)$$

where

$$k'_{UB_{n_r}} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \left(\frac{1}{\bar{d}_{ij}^2}\right)^{n_r} \leq \epsilon'_{n_r}(\text{snr}) \leq k'_{LB_{n_r}} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \left(\frac{1}{\bar{d}_{ij}^2}\right)^{n_r} \quad (82)$$

and  $k'_{LB_{n_r}}$  and  $k'_{UB_{n_r}}$  are given by (31) and (32), respectively.

*Proof:* The basis of the proof is the calculation of the probability density function of the squared pairwise Euclidean distance between two arbitrary (noiseless) receive vectors given by:

$$d_{ij}^2 = \text{tr}(\mathbf{H}_w \mathbf{\Delta}_{ij} \mathbf{H}_w^\dagger) \quad (83)$$

where  $\mathbf{\Delta}_{ij} = (\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^\dagger$ . Denote the eigenvalue decomposition of the positive semi-definite matrix  $\mathbf{\Delta}_{ij}$  by  $\mathbf{\Delta}_{ij} = \mathbf{W}_{ij} \mathbf{\Lambda}_{ij} \mathbf{W}_{ij}^\dagger$ , where  $\mathbf{W}_{ij}$  is a unitary matrix and  $\mathbf{\Lambda}_{ij} = \text{diag}(\lambda_{ij}, 0, \dots, 0)$  is a diagonal matrix with a single non-zero diagonal element  $\lambda_{ij} = \text{tr}(\mathbf{\Lambda}_{ij}) = \text{tr}(\mathbf{\Delta}_{ij}) = \text{tr}((\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^\dagger) = \|\mathbf{x}_i - \mathbf{x}_j\|^2 = \bar{d}_{ij}^2$  (note that there is a single non-zero diagonal element  $\lambda_{ij} > 0$  because the matrix  $\mathbf{\Delta}_{ij}$  is unit rank). It is possible to show (see [3]) that the distribution of

$$\text{tr}(\mathbf{H}_w \mathbf{\Delta}_{ij} \mathbf{H}_w^\dagger) = \text{tr}(\mathbf{H}_w \mathbf{W}_{ij} \mathbf{\Lambda}_{ij} \mathbf{W}_{ij}^\dagger \mathbf{H}_w^\dagger) \quad (84)$$

is equal to the distribution of

$$\text{tr}(\mathbf{H}_w \mathbf{\Lambda}_{ij} \mathbf{H}_w^\dagger) = \sum_{k=1}^{n_r} \lambda_{ij} |\xi_k|^2 \quad (85)$$

where  $\xi_k, k = 1, \dots, n_r$ , are independent circularly symmetric complex Gaussian random variables with zero-mean and unit-variance. This is a central chi-square (or gamma) distribution with  $2n_r$  degrees of freedom with probability density function given by [31]:

$$p_{d_{ij}^2}(d_{ij}^2) = \frac{1}{\lambda_{ij}^{n_r} (n_r - 1)!} \cdot (d_{ij}^2)^{n_r - 1} \cdot e^{-\frac{d_{ij}^2}{\lambda_{ij}}}, \quad d_{ij}^2 \geq 0 \quad (86)$$

The high-snr expansion of the constrained capacity follows immediately from Theorem 2, by noting that the function  $\frac{1}{\lambda_{ij}^{n_r} (n_r - 1)!} \cdot (d_{ij}^2)^{n_r - 1} \cdot e^{-\frac{d_{ij}^2}{\lambda_{ij}}}$  and its higher-order derivatives are continuous and integrable on  $[0, \infty)$ ,  $p_{d_{ij}^2}^{(k)}(0) = 0, k = 0, \dots, n_r - 2$ , and  $p_{d_{ij}^2}^{(n_r - 1)}(0) = \frac{1}{\lambda_{ij}^{n_r}} \neq 0$ . ■

Theorem 3 defines the impact of the number of transmit and receive antennas as well as the characteristics of the signalling scheme on the constrained capacity of the canonical i.i.d. Rayleigh fading coherent channel with arbitrary equiprobable discrete inputs, in the regime of high snr. In particular, in view of the bounds in (82), this Theorem is consistent with the fact that constellations with poor sphere-packing properties require more signal-to-noise ratio than constellation with good sphere-packing properties in order to achieve a certain target constrained capacity in the regime of high snr.

The number of receive antennas controls the rate at which the constrained capacity value tends to its infinite-snr value, so that a multiple-receive antenna system requires less signal-to-noise ratio than a single-receive antenna system in order to achieve a certain target constrained capacity in the regime of high snr. The number of transmit antennas, as opposed to its effect on the capacity of the canonical i.i.d. multiple-antenna Rayleigh fading coherent channel [3], does not affect directly the multiplexing ability of the system because the infinite-snr constrained capacity value depends solely on the number of constellation points or vectors. Instead, a higher-dimensional (complex) space, in comparison to a lower-dimensional one, enables the construction of more efficiently packed multi-dimensional constellations [33]. This, in view of the bounds in (82), suggests that a multiple-transmit antenna system requires potentially less signal-to-noise ratio than a single-transmit antenna system in order to achieve a certain target constrained capacity in the regime of high snr. Figures 8 and 9, which illustrate these aspects, also demonstrate that the asymptotic expansions capture this behavior of the constrained capacity in the regime of high-snr.

Theorem 3 also suggests that appropriate constellations for communication over the canonical i.i.d. multiple-antenna Rayleigh fading coherent channel – in the sense of optimization of the constrained capacity – optimize the quantity:

$$\sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \left( \frac{1}{\bar{d}_{ij}^2} \right)^{n_r} \approx K \cdot \left( \frac{1}{\bar{d}_{\min}^2} \right)^{n_r} \quad (87)$$

where  $\bar{d}_{\min}^2 = \min_{i \neq j} \bar{d}_{ij}^2$  represents the minimum squared Euclidean distance between pairs of constellation points or vectors and  $K$  represents the number of pairs of points or vectors with pairwise squared Euclidean distance equal to the minimum squared Euclidean distance, a quantity akin to the kissing number. Interestingly, a different compromise appears to exist between  $K$  and  $\bar{d}_{\min}^2$  in canonical i.i.d Rayleigh fading coherent channels where only

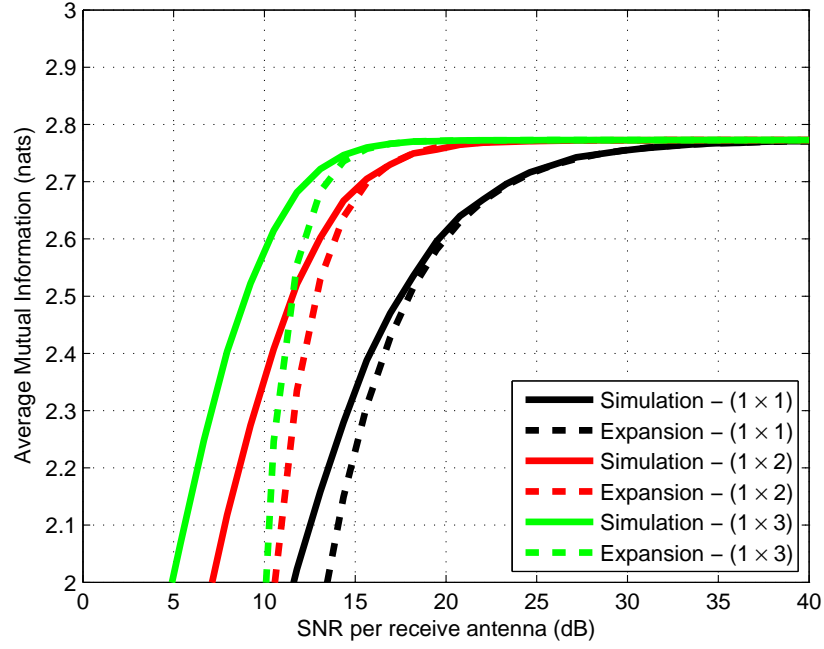


Fig. 8. Comparison of constrained capacity for  $1 \times 1$ ,  $1 \times 2$  and  $1 \times 3$  Rayleigh fading coherent channels with 16-QAM inputs.

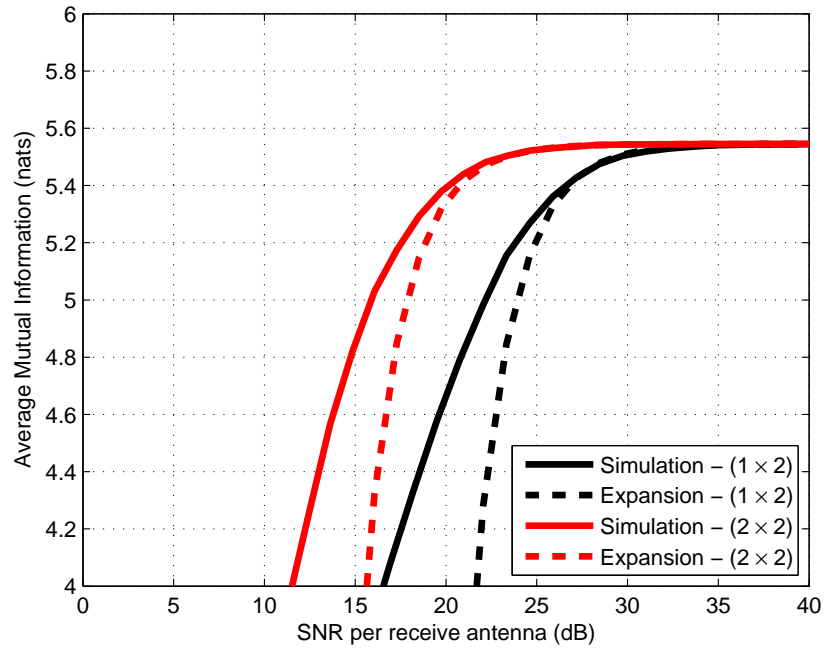


Fig. 9. Comparison of constrained capacity for a  $1 \times 2$  Rayleigh fading coherent channel with a 256-QAM input and a  $2 \times 2$  Rayleigh fading coherent channel with 16-QAM inputs.

the receiver knows the channel state than in fading channels where both the transmitter and the receiver know the channel state. Note that, when the channel matrix is known to the transmitter and the receiver, the high-snr expansion of the constrained capacity depends exponentially on  $\bar{d}_{\min}^2$  [16].

## V. OTHER FADING COHERENT CHANNELS

It is also relevant to characterize the constrained capacity of non-canonical fading coherent channels with arbitrary equiprobable discrete inputs in the regime of high snr. We study models that arise in many practical scenarios such as: *i*) the antenna-correlated Rayleigh fading coherent channel; and *ii*) the Ricean fading coherent channel.<sup>6</sup>

### A. Antenna-Correlated Rayleigh Fading Channels

We write  $\Delta_{ij} = (\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^\dagger$ . We denote the eigenvalue decomposition of the positive semi-definite matrix  $\Theta_{\mathbf{T}}^{\frac{1}{2}} \Delta_{ij} \Theta_{\mathbf{T}}^{\frac{1}{2}}$  by  $\Theta_{\mathbf{T}}^{\frac{1}{2}} \Delta_{ij} \Theta_{\mathbf{T}}^{\frac{1}{2}} = \mathbf{W}_{\mathbf{T}ij} \Lambda_{\mathbf{T}ij} \mathbf{W}_{\mathbf{T}ij}^\dagger$ , where  $\mathbf{W}_{\mathbf{T}ij}$  is a unitary matrix and  $\Lambda_{\mathbf{T}ij} = \text{diag}(\lambda_{T_{ij}}, 0, \dots, 0)$  is a diagonal matrix with at most a single non-zero diagonal element  $\lambda_{T_{ij}} \geq 0$  (note that there is at most a single non-zero diagonal element  $\lambda_{T_{ij}} \geq 0$  because the matrix  $\Theta_{\mathbf{T}}^{\frac{1}{2}} \Delta_{ij} \Theta_{\mathbf{T}}^{\frac{1}{2}}$  is at most unit rank). We also denote the eigenvalue decomposition of the positive semi-definite matrix  $\Theta_{\mathbf{R}}$  by  $\Theta_{\mathbf{R}} = \mathbf{W}_{\mathbf{R}} \Lambda_{\mathbf{R}} \mathbf{W}_{\mathbf{R}}^\dagger$ , where  $\mathbf{W}_{\mathbf{R}}$  is a unitary matrix and  $\Lambda_{\mathbf{R}} = \text{diag}(\lambda_{R_k})$  is a diagonal matrix with diagonal elements  $\lambda_{R_k} \geq 0, k = 1, \dots, n_r$ .

The following Theorem, which represents a generalization of Theorem 3, defines the high-snr behavior of the constrained capacity of antenna-correlated Rayleigh fading coherent channels. The Theorem concentrates on non-degenerate scenarios where  $\lambda_{T_{ij}} > 0, \forall i \neq j$ , and  $\lambda_{R_k} > 0, \forall k$ .<sup>7</sup>

*Theorem 4:* In the regime of high-snr, the constrained capacity of the  $n_r \times n_t$  antenna-correlated Rayleigh fading coherent channel with arbitrary equiprobable discrete inputs obeys:

$$\bar{I}(\text{snr}) = \log M - \epsilon'_{n_r}(\text{snr}) \cdot \frac{1}{\text{snr}^{n_r}} + \mathcal{O}\left(\frac{1}{\text{snr}^{n_r+1}}\right) \quad (88)$$

where

$$k'_{UB_{n_r}} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ i \neq j}}^M \left(\frac{1}{\lambda_{T_{ij}}}\right)^{n_r} \cdot \left(\prod_{k=1}^{n_r} \frac{1}{\lambda_{R_k}}\right) \leq \epsilon'_{n_r}(\text{snr}) \leq k'_{LB_{n_r}} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ i \neq j}}^M \left(\frac{1}{\lambda_{T_{ij}}}\right)^{n_r} \cdot \left(\prod_{k=1}^{n_r} \frac{1}{\lambda_{R_k}}\right) \quad (89)$$

and  $k'_{LB_{n_r}}$  and  $k'_{UB_{n_r}}$  are given by (31) and (32), respectively.

*Proof:* The squared pairwise Euclidean distance between two arbitrary (noiseless) receive vectors is given by:

$$d_{ij}^2 = \text{tr}\left(\Theta_{\mathbf{R}}^{\frac{1}{2}} \mathbf{H}_{\mathbf{w}} \Theta_{\mathbf{T}}^{\frac{1}{2}} \Delta_{ij} \Theta_{\mathbf{T}}^{\frac{1}{2}} \mathbf{H}_{\mathbf{w}}^\dagger \Theta_{\mathbf{R}}^{\frac{1}{2}}\right) \quad (90)$$

<sup>6</sup>We consider for simplicity exclusively Ricean fading channels with no transmit or receive antenna correlation.

<sup>7</sup>The Theorem also considers only scenarios where the eigenvalues  $\lambda_{R_k}, k = 1, \dots, n_r$ , are either all distinct or all equal (necessarily to one). The Theorem does not consider scenarios where there are groups of identical eigenvalues. This generalization requires considerable algebraic manipulation without adding much relevant insight. Note that  $\lambda_{R_1} = \lambda_{R_2} = \dots = \lambda_{R_{n_r}} = 1$  corresponds to a scenario where the receive antennas are uncorrelated.

It is possible to show (see [3]) that the distribution of

$$\text{tr}\left(\Theta_{\mathbf{R}}^{\frac{1}{2}}\mathbf{H}_{\mathbf{w}}\Theta_{\mathbf{T}}^{\frac{1}{2}}\Delta_{ij}\Theta_{\mathbf{T}}^{\frac{1}{2}}\mathbf{H}_{\mathbf{w}}^{\dagger}\Theta_{\mathbf{R}}^{\frac{1}{2}}\right) = \text{tr}\left(\mathbf{W}_{\mathbf{R}}\Lambda_{\mathbf{R}}^{\frac{1}{2}}\mathbf{W}_{\mathbf{R}}^{\dagger}\mathbf{H}_{\mathbf{w}}\mathbf{W}_{\mathbf{T}_{ij}}\Lambda_{\mathbf{T}_{ij}}\mathbf{W}_{\mathbf{T}_{ij}}^{\dagger}\mathbf{H}_{\mathbf{w}}^{\dagger}\mathbf{W}_{\mathbf{R}}\Lambda_{\mathbf{R}}^{\frac{1}{2}}\mathbf{W}_{\mathbf{R}}^{\dagger}\right) \quad (91)$$

is equal to the distribution of

$$\text{tr}\left(\Lambda_{\mathbf{R}}^{\frac{1}{2}}\mathbf{H}_{\mathbf{w}}\Lambda_{\mathbf{T}_{ij}}\mathbf{H}_{\mathbf{w}}^{\dagger}\Lambda_{\mathbf{R}}^{\frac{1}{2}}\right) = \sum_{k=1}^{n_r} \lambda_{T_{ij}} \lambda_{R_k} |\xi_k|^2 \quad (92)$$

where  $\xi_k, k = 1, \dots, n_r$ , are independent circularly symmetric complex Gaussian random variables with zero-mean and unit-variance.

Assume that the eigenvalues of the receive correlation matrix are all equal (necessarily to one), i.e.,  $\lambda_{R_1} = \lambda_{R_2} = \dots = \lambda_{R_{n_r}} = 1$ . Then, the probability density function of  $d_{ij}^2$  is given by [31]:

$$p_{d_{ij}^2}(d_{ij}^2) = \frac{1}{\lambda_{T_{ij}}^{n_r} (n_r - 1)!} \cdot (d_{ij}^2)^{n_r - 1} \cdot e^{-\frac{d_{ij}^2}{\lambda_{T_{ij}}}}, \quad d_{ij}^2 \geq 0 \quad (93)$$

The high-snr characterization of the constrained capacity follows from Theorem 2, by noting that

$$p_{d_{ij}^2}^{(k)}(0) = 0, \quad k = 0, \dots, n_r - 2 \quad (94)$$

and

$$p_{d_{ij}^2}^{(n_r - 1)}(0) = \frac{1}{\lambda_{T_{ij}}^{n_r}} \neq 0 \quad (95)$$

Assume now that the eigenvalues of the receive correlation matrix are all distinct, i.e.,  $\lambda_{R_1} \neq \lambda_{R_2} \neq \dots \neq \lambda_{R_{n_r}}$ . Then, the probability density function of  $d_{ij}^2$  is given by [34]:

$$p_{d_{ij}^2}(d_{ij}^2) = \sum_{k=1}^{n_r} \frac{1}{\lambda_{T_{ij}} \lambda_{R_k} \prod_{\substack{k'=1 \\ k' \neq k}}^{n_r} \left(1 - \frac{\lambda_{R_{k'}}}{\lambda_{R_k}}\right)} \cdot e^{-\frac{d_{ij}^2}{\lambda_{T_{ij}} \lambda_{R_k}}}, \quad d_{ij}^2 \geq 0 \quad (96)$$

The high-snr characterization of the constrained capacity also follows from Theorem 2, by noting that

$$p_{d_{ij}^2}^{(k)}(0) = 0, \quad k = 0, \dots, n_r - 2 \quad (97)$$

and

$$p_{d_{ij}^2}^{(n_r - 1)}(0) = \frac{1}{\lambda_{T_{ij}}^{n_r}} \cdot \prod_{k=1}^{n_r} \frac{1}{\lambda_{R_k}} \neq 0 \quad (98)$$

It is also simple to show that the conditions for the application of Theorem 2 are satisfied, i.e., the functions:

$$\frac{1}{\lambda_{T_{ij}}^{n_r} (n_r - 1)!} \cdot (d_{ij}^2)^{n_r - 1} \cdot e^{-\frac{d_{ij}^2}{\lambda_{T_{ij}}}} \quad (99)$$

and

$$\sum_{k=1}^{n_r} \frac{1}{\lambda_{T_{ij}} \lambda_{R_k} \prod_{\substack{k'=1 \\ k' \neq k}}^{n_r} \left(1 - \frac{\lambda_{R_{k'}}}{\lambda_{R_k}}\right)} \cdot e^{-\frac{d_{ij}^2}{\lambda_{T_{ij}} \lambda_{R_k}}} \quad (100)$$

and their higher-order derivatives are continuous and integrable on  $[0, \infty)$ . ■

It is also instructive to examine the implications of degenerate conditions where  $\lambda_{T_{ij}} = 0$  for some values of  $i$  and  $j$  and  $\lambda_{R_k} = 0$  for some values of  $k$ . Assume that the receive correlation matrix  $\Theta_{\mathbf{R}}$  has  $n'$  non-zero eigenvalues  $\lambda_{R_k} > 0$  and  $n'' = n_r - n'$  zero eigenvalues  $\lambda_{R_k} = 0$ , so that there are  $n''$  perfectly correlated paths. Assume also that some of the matrices  $\Theta_{\mathbf{T}}^{\frac{1}{2}} \Delta_{ij} \Theta_{\mathbf{T}}^{\frac{1}{2}}$  have eigenvalue  $\lambda_{T_{ij}} = 0$ , so that the vector  $\mathbf{x}_i - \mathbf{x}_j$  lies in the null space of the transmit correlation matrix  $\Theta_{\mathbf{T}}$ . This is also equivalent to saying that the receiver cannot distinguish some of the transmit vectors. By letting  $\log M' = H\left(\Theta_{\mathbf{R}}^{\frac{1}{2}} \mathbf{H}_{\mathbf{w}} \Theta_{\mathbf{T}}^{\frac{1}{2}} \mathbf{x} | \mathbf{H}_{\mathbf{w}}\right) < H(\mathbf{x}) = \log M$ , a simple modification of the Theorems reveals that as  $\text{snr} \rightarrow \infty$  the constrained capacity behaves as follows:

$$\bar{I}(\text{snr}) = \log M' - \epsilon'_{n'}(\text{snr}) \cdot \frac{1}{\text{snr}^{n'}} + \mathcal{O}\left(\frac{1}{\text{snr}^{n'+1}}\right) \quad (101)$$

where

$$k'_{UB_{n'}} \cdot \sum_{\lambda_{T_{ij}} > 0} \left(\frac{1}{\lambda_{T_{ij}}}\right)^{n'} \cdot \prod_{\lambda_{R_k} > 0} \frac{1}{\lambda_{R_k}} \leq \epsilon'_{n'}(\text{snr}) \leq k'_{LB_{n'}} \cdot \sum_{\lambda_{T_{ij}} > 0} \left(\frac{1}{\lambda_{T_{ij}}}\right)^{n'} \cdot \prod_{\lambda_{R_k} > 0} \frac{1}{\lambda_{R_k}} \quad (102)$$

and  $k'_{LB_{n'}}$  and  $k'_{UB_{n'}}$  are given by (31) and (32), respectively. This shows that degenerate scenarios affect the constrained capacity infinite-snr value as well as the rate at which the constrained capacity tends to the infinite-snr value.

In general, and the effect of degenerate conditions apart, the bounds in (89) suggest that the constrained capacity of a Rayleigh fading channel with transmit and receive antenna correlation is lower than the constrained capacity of a Rayleigh fading channel without antenna correlation for a certain signal-to-noise ratio in the regime of high snr. This is due to the fact that the value of the bounds in the correlated scenario in (89) is higher than the value of the bounds in the uncorrelated scenario in (82). For example, in the presence of receive antenna correlation, it is possible to show that

$$\prod_{k=1}^{n_r} \frac{1}{\lambda_{R_k}} \geq 1 \quad (103)$$

due to the fact that the function  $\prod_{k=1}^{n_r} \frac{1}{\lambda_{R_k}}$  is Schur-convex and the vector  $[\lambda_{R_1} \lambda_{R_2} \cdots \lambda_{R_{n_r}}]$  majorizes the vector  $[1 \ 1 \ \cdots \ 1]$  [35]. In the presence of transmit antenna correlation, it is possible to show that the convex function

$$\sum_{i=1}^M \sum_{\substack{j=1 \\ i \neq j}}^M \left(\frac{1}{\lambda_{T_{ij}}}\right)^{n_r} = \sum_{i=1}^M \sum_{\substack{j=1 \\ i \neq j}}^M \left(\frac{1}{\text{tr}((\mathbf{x}_i - \mathbf{x}_j)^{\dagger} \Theta_{\mathbf{T}} (\mathbf{x}_i - \mathbf{x}_j))}\right)^{n_r} \quad (104)$$

is minimized by  $\Theta_{\mathbf{T}} = \mathbf{I}$  in the set of unit-diagonal positive semi-definite matrices  $\Theta_{\mathbf{T}}$ . This result assumes that the constellations exhibit certain common symmetries (see also Section VI). Then,

$$\begin{aligned} \sum_{i=1}^M \sum_{\substack{j=1 \\ i \neq j}}^M \left(\frac{1}{\lambda_{T_{ij}}}\right)^{n_r} &= \sum_{i=1}^M \sum_{\substack{j=1 \\ i \neq j}}^M \left(\frac{1}{\text{tr}((\mathbf{x}_i - \mathbf{x}_j)^{\dagger} \Theta_{\mathbf{T}} (\mathbf{x}_i - \mathbf{x}_j))}\right)^{n_r} \\ &\geq \sum_{i=1}^M \sum_{\substack{j=1 \\ i \neq j}}^M \left(\frac{1}{\text{tr}((\mathbf{x}_i - \mathbf{x}_j)^{\dagger} (\mathbf{x}_i - \mathbf{x}_j))}\right)^{n_r} = \sum_{i=1}^M \sum_{\substack{j=1 \\ i \neq j}}^M \left(\frac{1}{d_{ij}^2}\right)^{n_r} \end{aligned} \quad (105)$$

This is portrayed in Figure 10.

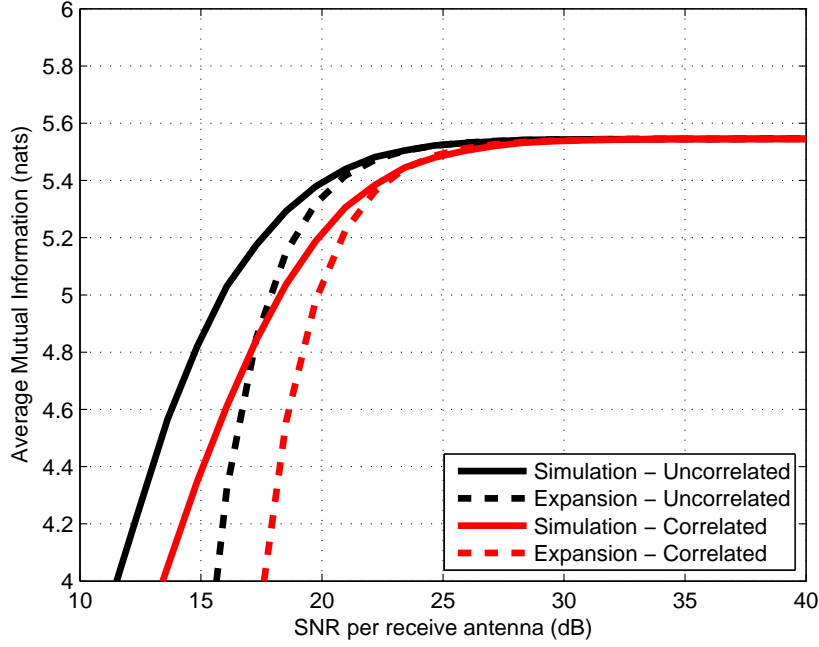


Fig. 10. Comparison of constrained capacity for a  $2 \times 2$  antenna-uncorrelated Rayleigh fading coherent channel with 16-QAM inputs and a  $2 \times 2$  antenna-correlated Rayleigh fading coherent channel with 16-QAM inputs ( $\Phi_{\mathbf{T}} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$  and  $\Phi_{\mathbf{R}} = \begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix}$ ).

### B. Ricean Fading Channels

We also write  $\Delta_{ij} = (\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^\dagger$ . We denote the eigenvalue decomposition of the positive semi-definite matrix  $\Delta_{ij}$  by  $\Delta_{ij} = \mathbf{W}_{ij} \Lambda_{ij} \mathbf{W}_{ij}^\dagger$ , where  $\mathbf{W}_{ij}$  is a unitary matrix and  $\Lambda_{ij} = \text{diag}(\lambda_{ij}, 0, \dots, 0)$  is a diagonal matrix with a single non-zero diagonal element  $\lambda_{ij} = \text{tr}(\Lambda_{ij}) = \text{tr}(\Delta_{ij}) = \text{tr}((\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^\dagger) = \|(\mathbf{x}_i - \mathbf{x}_j)\|^2 = \bar{d}_{ij}^2$ .

The following Theorem, which also represents a generalization of Theorem 3, defines the high-snr behavior of the constrained capacity of Ricean fading coherent channels.

*Theorem 5:* In the regime of high-snr, the constrained capacity of the  $n_r \times n_t$  Ricean fading coherent channel with arbitrary equiprobable discrete inputs obeys:

$$\bar{I}(\text{snr}) = \log M - \epsilon'_{n_r}(\text{snr}) \cdot \frac{1}{\text{snr}^{n_r}} + \mathcal{O}\left(\frac{1}{\text{snr}^{n_r+1}}\right) \quad (106)$$

where

$$\begin{aligned}
& k'_{UB_{n_r}} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \left( \frac{1}{\bar{d}_{ij}^2 \cdot \frac{1}{K+1}} \right)^{n_r} \cdot e^{-K \cdot \text{tr}(\mathbf{H}_0 \mathbf{W}_{ij} \mathbf{e}_1 \mathbf{e}_1^\dagger \mathbf{W}_{ij}^\dagger \mathbf{H}_0^\dagger)} \\
& \leq \epsilon'_{n_r}(\text{snr}) \leq \\
& k'_{LB_{n_r}} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \left( \frac{1}{\bar{d}_{ij}^2 \cdot \frac{1}{K+1}} \right)^{n_r} \cdot e^{-K \cdot \text{tr}(\mathbf{H}_0 \mathbf{W}_{ij} \mathbf{e}_1 \mathbf{e}_1^\dagger \mathbf{W}_{ij}^\dagger \mathbf{H}_0^\dagger)}
\end{aligned} \tag{107}$$

and  $k'_{LB_{n_r}}$  and  $k'_{UB_{n_r}}$  are given by (31) and (32), respectively.

*Proof:* The squared pairwise Euclidean distance  $\bar{d}_{ij}^2$  between two arbitrary (noiseless) receive vectors is given by:

$$\bar{d}_{ij}^2 = \text{tr} \left( \left( \sqrt{\frac{K}{K+1}} \cdot \mathbf{H}_0 + \sqrt{\frac{1}{K+1}} \cdot \mathbf{H}_w \right) \Delta_{ij} \left( \sqrt{\frac{K}{K+1}} \cdot \mathbf{H}_0 + \sqrt{\frac{1}{K+1}} \cdot \mathbf{H}_w \right)^\dagger \right) \tag{108}$$

It is also possible to show (see [3]) that the distribution of

$$\begin{aligned}
& \text{tr} \left( \left( \sqrt{\frac{K}{K+1}} \cdot \mathbf{H}_0 + \sqrt{\frac{1}{K+1}} \cdot \mathbf{H}_w \right) \Delta_{ij} \left( \sqrt{\frac{K}{K+1}} \cdot \mathbf{H}_0 + \sqrt{\frac{1}{K+1}} \cdot \mathbf{H}_w \right)^\dagger \right) = \\
& = \text{tr} \left( \left( \sqrt{\frac{K}{K+1}} \cdot \mathbf{H}_0 + \sqrt{\frac{1}{K+1}} \cdot \mathbf{H}_w \right) \mathbf{W}_{ij} \Lambda_{ij} \mathbf{W}_{ij}^\dagger \left( \sqrt{\frac{K}{K+1}} \cdot \mathbf{H}_0 + \sqrt{\frac{1}{K+1}} \cdot \mathbf{H}_w \right)^\dagger \right)
\end{aligned} \tag{109}$$

is equal to the distribution of

$$\begin{aligned}
& \text{tr} \left( \left( \sqrt{\frac{K}{K+1}} \cdot \mathbf{W}_{ij}^\dagger \mathbf{H}_0 \mathbf{W}_{ij} + \sqrt{\frac{1}{K+1}} \cdot \mathbf{H}_w \right) \Lambda_{ij} \left( \sqrt{\frac{K}{K+1}} \cdot \mathbf{W}_{ij}^\dagger \mathbf{H}_0 \mathbf{W}_{ij} + \sqrt{\frac{1}{K+1}} \cdot \mathbf{H}_w \right)^\dagger \right) = \\
& = \sum_{k=1}^{n_r} \lambda_{ij} |\xi_k|^2
\end{aligned} \tag{110}$$

where  $\xi_k, k = 1, \dots, n_r$ , are independent circularly symmetric complex Gaussian random variables with mean  $\sqrt{\frac{K}{K+1}} \cdot \mathbf{e}_k^\dagger \mathbf{W}_{ij}^\dagger \mathbf{H}_0 \mathbf{W}_{ij} \mathbf{e}_1$  and variance  $\frac{1}{K+1}$ , respectively. This is a non-central chi-square (or gamma) distribution with  $2n_r$  degrees of freedom with probability density function given by [31]:

$$p_{\bar{d}_{ij}^2}(\bar{d}_{ij}^2) = \frac{1}{\lambda_{ij} \cdot \frac{1}{K+1}} \cdot \left( \frac{\bar{d}_{ij}^2}{s^2} \right)^{\frac{n_r-1}{2}} \cdot e^{-\frac{s^2 + \bar{d}_{ij}^2}{\lambda_{ij} \cdot \frac{1}{K+1}}} \cdot I_{n_r-1} \left( \frac{2\sqrt{s^2 \bar{d}_{ij}^2}}{\lambda_{ij} \cdot \frac{1}{K+1}} \right), \quad \bar{d}_{ij}^2 \geq 0 \tag{111}$$

where  $s^2 = \lambda_{ij} \cdot \frac{K}{K+1} \cdot \left( \sum_{k=1}^{n_r} \left\| \mathbf{e}_k^\dagger \mathbf{W}_{ij}^\dagger \mathbf{H}_0 \mathbf{W}_{ij} \mathbf{e}_1 \right\|^2 \right) = \lambda_{ij} \cdot \frac{K}{K+1} \cdot \text{tr}(\mathbf{H}_0 \mathbf{W}_{ij} \mathbf{e}_1 \mathbf{e}_1^\dagger \mathbf{W}_{ij}^\dagger \mathbf{H}_0^\dagger)$  and  $I_n(\cdot)$  is the  $n$ th-order modified Bessel function of the first-kind given by:

$$I_n(x) = \sum_{k=0}^{\infty} \frac{1}{k! \Gamma(n+k+1)} \cdot \left( \frac{x}{2} \right)^{n+2k}, \quad x \geq 0 \tag{112}$$

The high-snr characterization of the constrained capacity also follows from Theorem 2, by noting that

$$p_{\bar{d}_{ij}^2}^{(k)}(0) = 0, \quad k = 0, \dots, n_r - 2 \tag{113}$$



and

$$p_{d_{ij}^2}^{(n_r-1)}(0) = \frac{1}{\left(\lambda_{ij} \cdot \frac{1}{K+1}\right)^{n_r}} \cdot e^{-\frac{s^2}{\lambda_{ij} \cdot \frac{1}{K+1}}} \neq 0 \quad (114)$$

It is also simple to show that the conditions for the application of Theorem 2 are satisfied, i.e., the function:

$$\frac{1}{\lambda_{ij} \cdot \frac{1}{K+1}} \cdot \left(\frac{d_{ij}^2}{s^2}\right)^{\frac{n_r-1}{2}} \cdot e^{-\frac{s^2+d_{ij}^2}{\lambda_{ij} \cdot \frac{1}{K+1}}} \cdot I_{n_r-1} \left( \frac{2\sqrt{s^2 d_{ij}^2}}{\lambda_{ij} \cdot \frac{1}{K+1}} \right) \quad (115)$$

and its higher-order derivatives are continuous and integrable on  $[0, \infty)$ .  $\blacksquare$

In general, the bounds in (107) suggest that the constrained capacity of a Ricean fading channel is normally higher than the constrained capacity of the canonical i.i.d. Rayleigh fading channel for a certain signal-to-noise ratio in the regime of high snr. This is due to the fact that the function

$$\begin{aligned} & \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \left( \frac{1}{\bar{d}_{ij}^2 \cdot \frac{1}{K+1}} \right)^{n_r} \cdot e^{-K \cdot \text{tr}(\mathbf{H}_0 \mathbf{W}_{ij} \mathbf{e}_1 \mathbf{e}_1^\dagger \mathbf{W}_{ij}^\dagger \mathbf{H}_0^\dagger)} = \\ & \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \left( \frac{1}{\bar{d}_{ij}^2 \cdot \frac{1}{K+1}} \right)^{n_r} \cdot e^{-K \cdot \text{tr}(\mathbf{a}_R \mathbf{a}_R^\dagger \mathbf{W}_{ij} \mathbf{e}_1 \mathbf{e}_1^\dagger \mathbf{W}_{ij}^\dagger \mathbf{a}_T \mathbf{a}_T^\dagger)} = \\ & \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \left( \frac{1}{\bar{d}_{ij}^2 \cdot \frac{1}{K+1}} \right)^{n_r} \cdot e^{-n_r \cdot K \cdot \text{tr}(\mathbf{a}_T^\dagger \mathbf{W}_{ij} \mathbf{e}_1 \mathbf{e}_1^\dagger \mathbf{W}_{ij}^\dagger \mathbf{a}_T)} \end{aligned} \quad (116)$$

is monotonically decreasing in the  $K$ -factor  $K$  provided that  $\text{tr}(\mathbf{a}_T^\dagger \mathbf{W}_{ij} \mathbf{e}_1 \mathbf{e}_1^\dagger \mathbf{W}_{ij}^\dagger \mathbf{a}_T) \geq 1, \forall i \neq j$ . This is portrayed in Figure 11.

However, as shown in Figure 12, and in contrast to the capacity of a multiple-antenna Ricean fading channel [12], the constrained capacity of a Ricean fading channel can also be lower than the constrained capacity of a Rayleigh fading channel for a certain signal-to-noise ratio in the regime of high snr. This aspect is illustrated further by the following toy-examples.

*Example 1:* Consider a  $2 \times 2$  Ricean fading channel with two equiprobable discrete inputs  $\mathbf{x}_1 = [1 \ 0]^\top$  and  $\mathbf{x}_2 = [0 \ 1]^\top$ , where  $\mathbf{H} = \sqrt{\frac{K}{K+1}} \cdot \mathbf{H}_0 + \sqrt{\frac{1}{K+1}} \cdot \mathbf{H}_w$  and  $\mathbf{H}_0 = \mathbf{a}_R \mathbf{a}_T^\dagger$  with  $\mathbf{a}_R = [1 \ 1]^\top$  and  $\mathbf{a}_T = [1 \ 1]^\top$ . Then,

$$k'_{UB_2} \cdot \frac{1}{2} \cdot (K+1)^2 \leq \epsilon'_2(\text{snr}) \leq k'_{LB_2} \cdot \frac{1}{2} \cdot (K+1)^2 \quad (117)$$

with  $k'_{LB_2}$  and  $k'_{UB_2}$  given by (31) and (32), respectively.

*Example 2:* Consider a  $2 \times 2$  Rayleigh fading channel with the two equiprobable discrete inputs  $\mathbf{x}_1 = [1 \ 0]^\top$  and  $\mathbf{x}_2 = [0 \ 1]^\top$ , where  $\mathbf{H} = \mathbf{H}_w$ . Then,

$$k'_{UB_2} \cdot \frac{1}{2} \leq \epsilon'_2(\text{snr}) \leq k'_{LB_2} \cdot \frac{1}{2} \quad (118)$$

with  $k'_{LB_2}$  and  $k'_{UB_2}$  given by (31) and (32), respectively.

Examples 1 and 2 suggest that, due to the bounds in (117) and (118), this Ricean fading channel needs a higher signal-to-noise ratio than the Rayleigh fading channel for  $K > 0$ , in order to achieve a certain target constrained

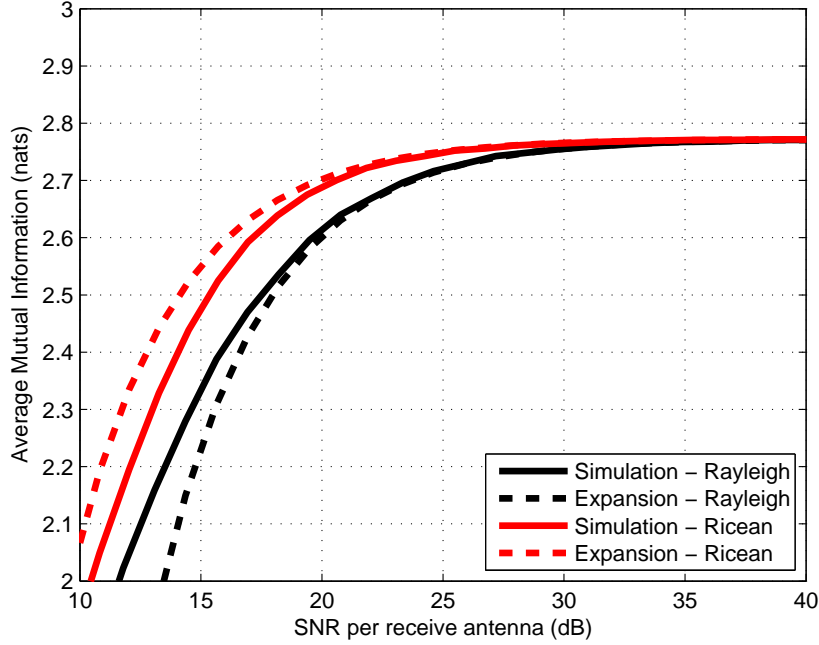


Fig. 11. Comparison of constrained capacity for a  $1 \times 1$  Rayleigh fading coherent channel with a 16-QAM input and a  $1 \times 1$  Ricean fading coherent channel with a 16-QAM input ( $K = 2$ ,  $\mathbf{H}_0 = \mathbf{1}$ ).

capacity in the regime of high snr. Figure 13, which shows only the simulated constrained capacity, confirms this behavior. This is due to the fact that one cannot rely on the deterministic component but rather, in comparison to a canonical i.i.d. Rayleigh fading channel, on the lower power random component in order to distinguish the transmit vectors. Note that the transmit vector differences  $\mathbf{x}_1 - \mathbf{x}_2 = [1 \ -1]^T$  and  $\mathbf{x}_2 - \mathbf{x}_1 = [-1 \ 1]^T$  are orthogonal to the transmit array response  $\mathbf{a}_T = [1 \ 1]^T$ .

## VI. DESIGNS

The focus now is on the design of schemes for communication over multiple-antenna fading coherent channels driven by equiprobable discrete inputs, in the regime of high snr. In particular, we conceive designs for: *i*) optimal power allocation over a bank of parallel independent fading coherent channels; *ii*) optimal precoding for multiple-antenna fading coherent channels; and *iii*) space-time coding for the multiple-antenna canonical i.i.d. Rayleigh fading coherent channel. The design principle is based on the optimization of the lower bound to the asymptotic expansion of the constrained capacity, rather than the exact asymptotic expansion of the constrained capacity. It is interesting to note though that the procedure leads to very sharp designs. It is also interesting to note that some of the design criteria coincide with standard design criteria in the literature, most notably the space-time coding criteria, which have been derived based on other principles.

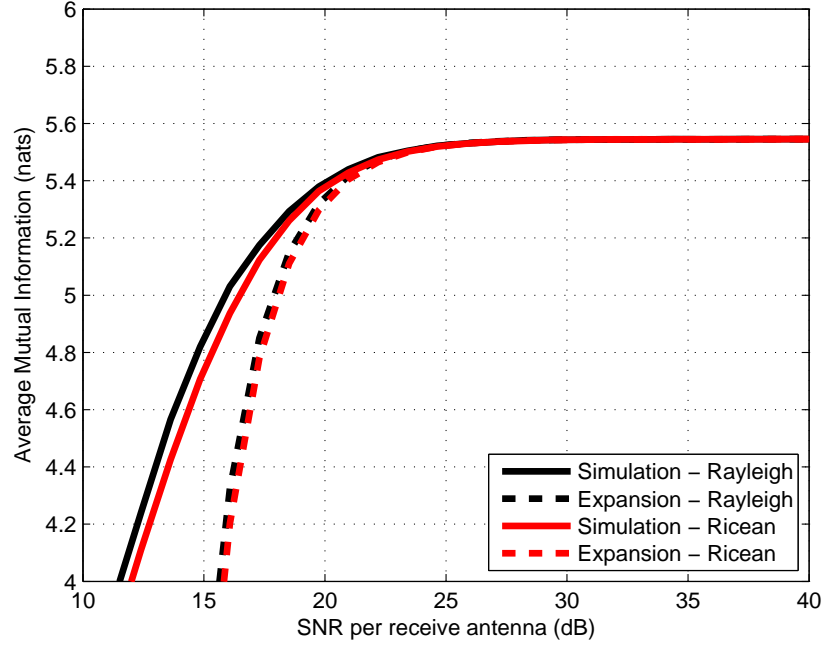


Fig. 12. Comparison of constrained capacity for a  $2 \times 2$  Rayleigh fading coherent channel with 16-QAM inputs and a  $2 \times 2$  Ricean fading coherent channel with 16-QAM inputs ( $K = 3, \mathbf{H}_0 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ ).

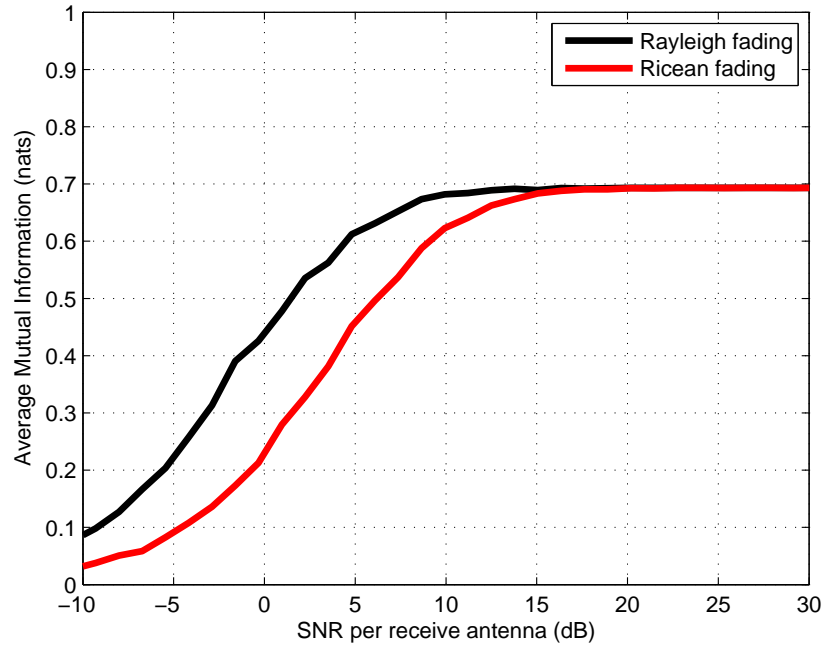


Fig. 13. Comparison of constrained capacity for the fading coherent channels in Examples 1 and 2 ( $K = 2$ ).

### A. Power Allocation in a Bank of Parallel Independent Fading Channels

We consider a bank of  $K$  parallel independent fading channels which can be modeled as follows:

$$y_k = \sqrt{\text{snr}} \cdot h_k \sqrt{p_k} x_k + n_k, \quad k = 1, \dots, K \quad (119)$$

for a single channel use, where  $y_k \in \mathbb{C}$  is the  $k$ th sub-channel complex receive symbol,  $x_k \in \mathbb{C}$  is the  $k$ th sub-channel complex transmit symbol,  $h_k \in \mathbb{C}$  is the  $k$ th sub-channel random complex fading coefficient and  $n_k \in \mathbb{C}$  is a circularly symmetric complex Gaussian noise random variable with zero-mean and unit-variance. The variable  $p_k \in \mathbb{R}_0^+$  corresponds to the power injected into the  $k$ th sub-channel. The scaling factor  $\text{snr}$  relates to the signal-to-noise ratio. We assume that the transmit symbols  $x_k, k = 1, \dots, K$ , conform to equiprobable unit-power discrete constellations with cardinality  $M_k, k = 1, \dots, K$ . We also assume that all the random variables are independent.

The constrained capacity of the bank of parallel independent fading channels is given by:

$$\bar{I}(\text{snr}) = \sum_{k=1}^K \bar{I}_k(\text{snr}) = \sum_{k=1}^K \mathbb{E}_{h_k} \{I(x_k; \sqrt{\text{snr}} \cdot h_k \sqrt{p_k} x_k + n_k | h_k)\} \quad (120)$$

The objective is to determine the power allocation procedure that maximizes a lower bound to the constrained capacity in the asymptotic regime of high  $\text{snr}$ , subject to the total power constraint:

$$\sum_{k=1}^K p_k \leq P \quad (121)$$

with  $p_k \geq 0, k = 1, \dots, K$ .

Next, we consider a bank of parallel independent channels subject to either Rayleigh fading or Ricean fading. We denote the squared pairwise Euclidean distance between two distinct transmit symbols in a particular sub-channel  $k, x_{k_i}$  and  $x_{k_j}$ , by:

$$\bar{d}_{k_{ij}}^2 = \|x_{k_i} - x_{k_j}\|^2 \quad (122)$$

We also denote the squared pairwise Euclidean distance between two distinct noiseless receive symbols in a particular sub-channel  $k, h_k \sqrt{p_k} x_{k_i}$  and  $h_k \sqrt{p_k} x_{k_j}$ , by:

$$d_{k_{ij}}^2 = \|h_k \sqrt{p_k} (x_{k_i} - x_{k_j})\|^2 = p_k \|h_k\|^2 \|x_{k_i} - x_{k_j}\|^2 = p_k |h_k|^2 \bar{d}_{k_{ij}}^2 \quad (123)$$

1) *Rayleigh fading case:* In the Rayleigh fading case, the complex fading coefficients  $h_k \sim \mathcal{CN}(0, \sigma_{h_k}^2), k = 1, \dots, K$ . Therefore, the probability density function of the squared pairwise Euclidean distance between two distinct noiseless receive symbols in a particular sub-channel  $k$  is given by:

$$p_{d_{k_{ij}}^2}(d_{k_{ij}}^2) = \frac{1}{\bar{d}_{k_{ij}}^2 \sigma_{h_k}^2 p_k} \cdot e^{-\frac{d_{k_{ij}}^2}{\bar{d}_{k_{ij}}^2 \sigma_{h_k}^2 p_k}}, \quad d_{k_{ij}}^2 \geq 0 \quad (124)$$

The asymptotic characterization of the constrained capacity, which can be found by using Theorem 2 together with the properties of (124), is given by:

$$\bar{I}(\text{snr}) = \log M - \epsilon'_1(\text{snr}) \cdot \frac{1}{\text{snr}} + \mathcal{O}\left(\frac{1}{\text{snr}^2}\right) \quad (125)$$

where  $M = \prod_{k=1}^K M_k$  and

$$\sum_{k=1}^K k'_{UB_{1_k}} \sum_{i=1}^{M_k} \sum_{\substack{j=1 \\ j \neq i}}^{M_k} \frac{1}{\bar{d}_{kij}^2 \sigma_{h_k}^2 p_k} \leq \epsilon'_1(\text{snr}) \leq \sum_{k=1}^K k'_{LB_{1_k}} \sum_{i=1}^{M_k} \sum_{\substack{j=1 \\ j \neq i}}^{M_k} \frac{1}{\bar{d}_{kij}^2 \sigma_{h_k}^2 p_k} \quad (126)$$

and

$$k'_{UB_{1_k}} = \frac{1}{2M_k(M_k - 1)} \cdot \frac{1}{\sqrt{\pi}} \cdot 4^n \cdot \frac{\Gamma(5/2)}{\Gamma(3)} \quad (127)$$

$$k'_{LB_{1_k}} = \frac{2}{M_k} \cdot \frac{1}{\sqrt{\pi}} \cdot 4^n \cdot \frac{\Gamma(5/2)}{\Gamma(3)} \quad (128)$$

As  $\text{snr} \rightarrow \infty$ , it follows from (125) and the upper bound to  $\epsilon'_1(\text{snr})$  in (126) that the power allocation procedure that maximizes the lower bound to the constrained capacity is given by:

$$p_k^* = \alpha \cdot \frac{1}{\sqrt{\frac{1}{M_k} \sum_{i=1}^{M_k} \sum_{\substack{j=1 \\ j \neq i}}^{M_k} \bar{d}_{kij}^2 \sigma_{h_k}^2}} + o(1) \quad (129)$$

with

$$\alpha^{-1} = \frac{1}{P} \cdot \sum_{k=1}^K \frac{1}{\sqrt{\frac{1}{M_k} \sum_{i=1}^{M_k} \sum_{\substack{j=1 \\ j \neq i}}^{M_k} \bar{d}_{kij}^2 \sigma_{h_k}^2}} \quad (130)$$

Note that, given equal sub-channel constellations, the higher the average sub-channel strength (i.e., the higher  $\sigma_{h_k}^2$ ), then the lower the allocated power. Note also that the power allocation policy embodied in (129) and (130) in fact corresponds to the power allocation policy put forth in [24].

Figures 14 and 15 compare analysis to simulation for a bank of two parallel independent Rayleigh fading coherent channels driven by 16-QAM inputs. We observe that the optimal power allocation policy, obtained via direct optimization of the Monte-Carlo simulated constrained capacity, tends with the increase of the signal-to-noise ratio to the high-snr power allocation policy embodied in (129) and (130). We also observe that the constrained capacity associated with the optimal power allocation policy tends with the increase of the signal-to-noise ratio to the constrained capacity associated with the high-snr power allocation policy. It is important to note though that the designs, which are shown to be very good in the regime of high snr in Figures 14 and 15, are not sharp for parallel independent channels driven by distinct inputs. This is due to the fact that one can only bound rather than compute the exact value of the quantity akin to the MMSE dimension, which differs for different inputs.

2) *Ricean fading case:* In the Ricean fading case, the complex fading coefficients  $h_k \sim \mathcal{CN}(\mu_{h_k}, \sigma_{h_k}^2)$ ,  $k = 1, \dots, K$ . Therefore, the probability density function of the squared pairwise Euclidean distance between two distinct noiseless receive symbols in a particular sub-channel  $k$  is given by:

$$p_{d_{kij}^2}(d_{kij}^2) = \frac{1}{\bar{d}_{kij}^2 \sigma_{h_k}^2 p_k} \cdot e^{-\frac{s_k^2 + d_{kij}^2}{\bar{d}_{kij}^2 \sigma_{h_k}^2 p_k}} \cdot I_0\left(\frac{2\sqrt{s_k^2 d_{kij}^2}}{\bar{d}_{kij}^2 \sigma_{h_k}^2 p_k}\right), \quad d_{kij}^2 \geq 0 \quad (131)$$

where  $s_k^2 = \bar{d}_{kij}^2 \cdot p_k \cdot (\Re\{\mu_{h_k}\}^2 + \Im\{\mu_{h_k}\}^2)$  and

$$I_0(x) = \sum_{k=0}^{\infty} \frac{1}{k! \Gamma(k+1)} \cdot \left(\frac{x}{2}\right)^{2k}, \quad x \geq 0 \quad (132)$$

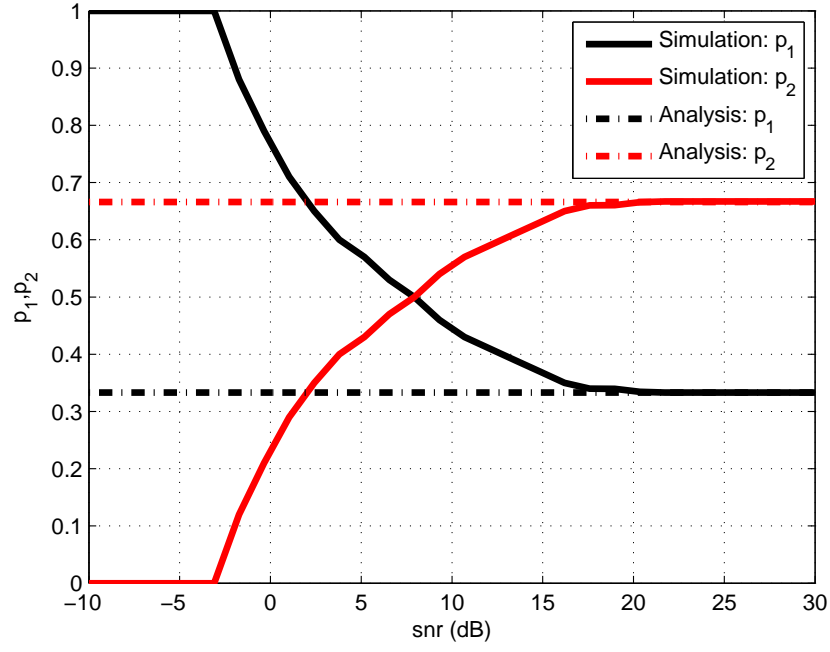


Fig. 14. Optimal power allocation in a bank of two parallel independent Rayleigh fading coherent channels driven by 16-QAM inputs ( $\sigma_{h_1}^2 = 4$  and  $\sigma_{h_2}^2 = 1$ ).

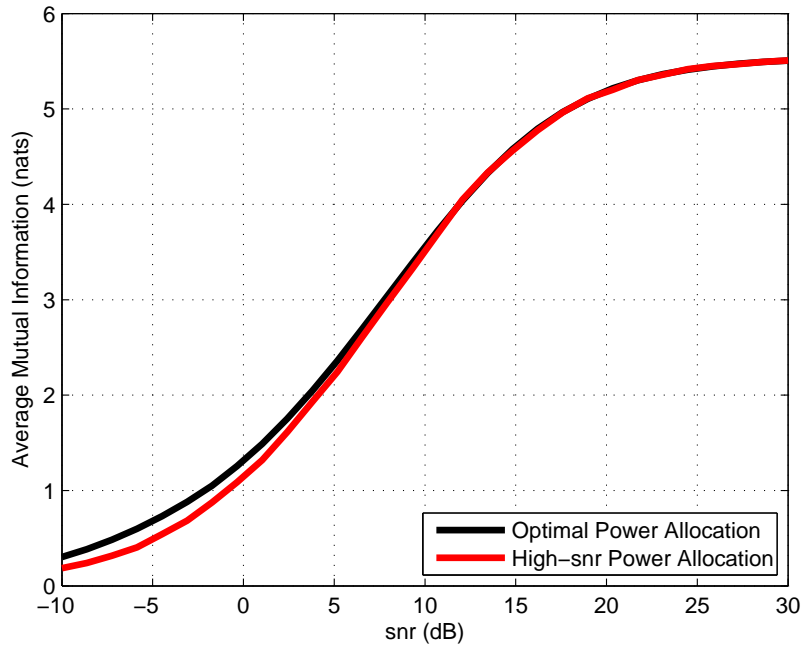


Fig. 15. Average mutual information in a bank of two parallel independent Rayleigh fading coherent channels driven by 16-QAM inputs ( $\sigma_{h_1}^2 = 4$  and  $\sigma_{h_2}^2 = 1$ ).

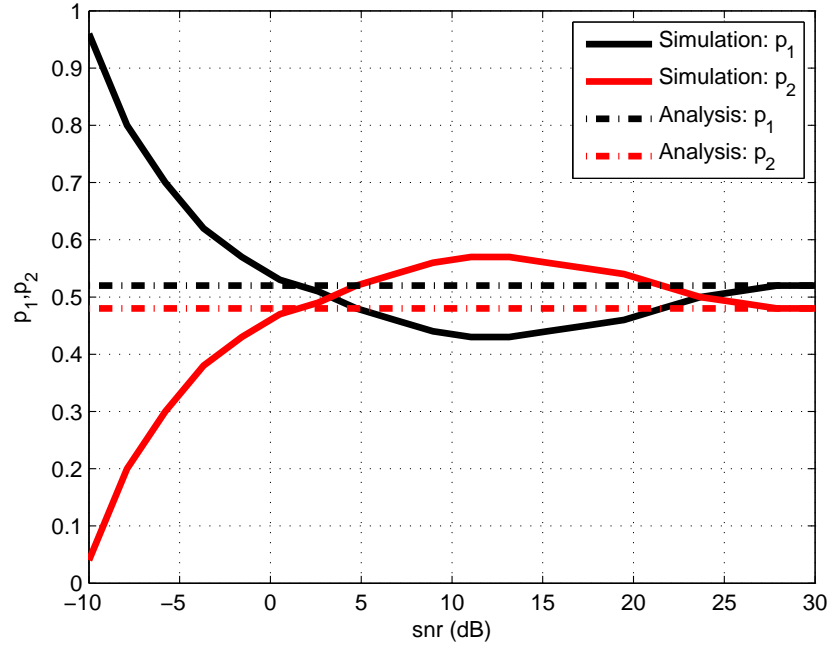


Fig. 16. Optimal power allocation in a bank of two parallel independent Ricean fading coherent channels driven by 16-QAM inputs ( $\mu_{h_1} = 1+j$ ,  $\mu_{h_2} = 1+j$ ,  $\sigma_{h_1}^2 = 4$  and  $\sigma_{h_2}^2 = 1$ ).

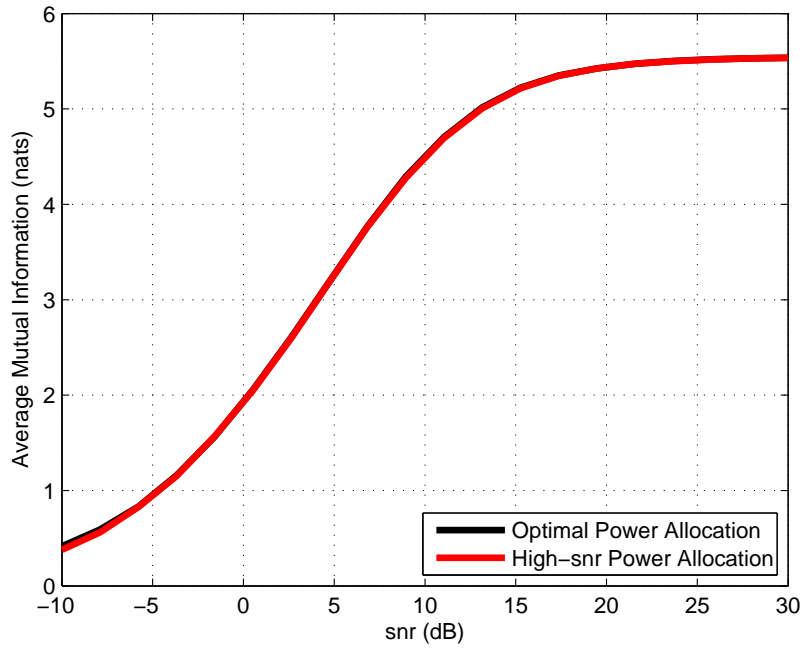


Fig. 17. Average mutual information in a bank of two parallel independent Ricean fading coherent channels driven by 16-QAM inputs ( $\mu_{h_1} = 1+j$ ,  $\mu_{h_2} = 1+j$ ,  $\sigma_{h_1}^2 = 4$  and  $\sigma_{h_2}^2 = 1$ ).

The asymptotic characterization of the constrained capacity, which can also be found by using Theorem 2 together with the properties of (131), is given by:

$$\bar{I}(\text{snr}) = \log M - \epsilon'_1(\text{snr}) \cdot \frac{1}{\text{snr}} + \mathcal{O}\left(\frac{1}{\text{snr}^2}\right) \quad (133)$$

where  $M = \prod_{k=1}^K M_k$  and

$$\begin{aligned} & \sum_{k=1}^K k'_{\text{UB}_{1k}} \sum_{i=1}^{M_k} \sum_{\substack{j=1 \\ j \neq i}}^{M_k} \frac{1}{\bar{d}_{kij}^2 \sigma_{h_k}^2 p_k} \cdot e^{-\frac{\Re\{\mu_{h_k}\}^2 + \Im\{\mu_{h_k}\}^2}{\sigma_{h_k}^2}} \\ & \leq \epsilon'_1(\text{snr}) \leq \\ & \sum_{k=1}^K k'_{\text{LB}_{1k}} \sum_{i=1}^{M_k} \sum_{\substack{j=1 \\ j \neq i}}^{M_k} \frac{1}{\bar{d}_{kij}^2 \sigma_{h_k}^2 p_k} \cdot e^{-\frac{\Re\{\mu_{h_k}\}^2 + \Im\{\mu_{h_k}\}^2}{\sigma_{h_k}^2}} \end{aligned} \quad (134)$$

and

$$k'_{\text{UB}_{1k}} = \frac{1}{2M_k(M_k - 1)} \cdot \frac{1}{\sqrt{\pi}} \cdot 4^n \cdot \frac{\Gamma(5/2)}{\Gamma(3)} \quad (135)$$

$$k'_{\text{LB}_{1k}} = \frac{2}{M_k} \cdot \frac{1}{\sqrt{\pi}} \cdot 4^n \cdot \frac{\Gamma(5/2)}{\Gamma(3)} \quad (136)$$

As  $\text{snr} \rightarrow \infty$ , it follows from (133) and the upper bound to  $\epsilon'_1(\text{snr})$  in (134) that the power allocation procedure that maximizes the lower bound to the constrained capacity is given by:

$$p_k^* = \alpha \cdot \frac{e^{-\frac{\Re\{\mu_{h_k}\}^2 + \Im\{\mu_{h_k}\}^2}{2\sigma_{h_k}^2}}}{\sqrt{\frac{1}{M_k} \sum_{i=1}^{M_k} \sum_{\substack{j=1 \\ j \neq i}}^{M_k} \bar{d}_{kij}^2 \sigma_{h_k}^2}} + o(1) \quad (137)$$

with

$$\alpha^{-1} = \frac{1}{P} \cdot \sum_{k=1}^K \frac{e^{-\frac{\Re\{\mu_{h_k}\}^2 + \Im\{\mu_{h_k}\}^2}{2\sigma_{h_k}^2}}}{\sqrt{\frac{1}{M_k} \sum_{i=1}^{M_k} \sum_{\substack{j=1 \\ j \neq i}}^{M_k} \bar{d}_{kij}^2 \sigma_{h_k}^2}} \quad (138)$$

Note now that the Ricean K-factor  $(\Re\{\mu_{h_k}\}^2 + \Im\{\mu_{h_k}\}^2)/\sigma_{h_k}^2$  has a direct impact on the power allocation procedure (compare (129) and (130) to (137) and (138)).

Figures 16 and 17, which compare analysis to simulation for a bank of two parallel independent Ricean fading coherent channels driven by 16-QAM inputs, also demonstrate that the optimal designs tend with the increase of the signal-to-noise ratio to the high-snr designs.<sup>8</sup> It is also important to note that the designs, which are shown to be very good in the regime of high snr in Figures 16 and 17, are also not sharp for parallel independent channels driven by distinct inputs, as discussed previously.

<sup>8</sup>Note that a more substantial difference between the constrained capacity associated with the optimal power allocation and the constrained capacity associated with the high-snr power allocation is expected for a bank of parallel independent fading coherent channels with more than two sub-channels. The simulations, however, would be very time consuming.



### B. Precoding in a Multiple-Antenna Fading Channel

We consider a linearly precoded multiple-antenna fading channel which can be modeled as follows:

$$\mathbf{y} = \sqrt{\text{snr}} \cdot \mathbf{H}\mathbf{P}\mathbf{x} + \mathbf{n} \quad (139)$$

for a single channel use, where  $\mathbf{y} \in \mathbb{C}^{n_r}$  is the vector of complex receive symbols,  $\mathbf{x} \in \mathbb{C}^{n_t}$  is the vector of complex transmit symbols,  $\mathbf{H} \in \mathbb{C}^{n_r \times n_t}$  is the random channel fading matrix (with  $\mathbb{E}\{\text{tr}(\mathbf{H}\mathbf{H}^\dagger)\} = n_t n_r$ ), and  $\mathbf{n} \in \mathbb{C}^{n_r}$  is a vector of independent circularly symmetric complex Gaussian noise random variables with zero-mean and unit-variance. The matrix  $\mathbf{P} \in \mathbb{C}^{n_t \times n_t}$  represents a linear precoder.<sup>9</sup> The scaling factor  $\text{snr}$  relates to the signal-to-noise ratio. We assume that the transmit vector conforms to an equiprobable multi-dimensional constellation with cardinality  $M$ , i.e.,  $\mathbf{x} \in \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M\}$  and  $\Pr(\mathbf{x}_1) = \Pr(\mathbf{x}_2) = \dots = \Pr(\mathbf{x}_M) = \frac{1}{M}$ , with  $\Sigma_{\mathbf{x}} = \mathbb{E}\{\mathbf{x}\mathbf{x}^\dagger\} = \frac{1}{n_t} \cdot \mathbf{I}$ . We also assume that all the random variables are independent.

The objective is to determine the precoder that maximizes a lower bound to the constrained capacity in the asymptotic regime of high  $\text{snr}$ , subject to the total power constraint:

$$\text{tr}(\mathbf{P}\mathbf{P}^\dagger) \leq P \quad (140)$$

We illustrate the design procedure for the canonical i.i.d. Rayleigh fading coherent channel and the antenna-correlated Rayleigh fading coherent channel. In particular, the availability of closed-form expressions for the bounds to the asymptotic expansions of the constrained capacity, which embody the effect of the precoder, leads to simple design methods based on numerical procedures – and occasionally analytic results – rather than time-consuming Monte Carlo simulation procedures.

1) *The Canonical i.i.d. Rayleigh Fading Channel:* In this channel model, where  $\mathbf{H} = \mathbf{H}_{\mathbf{w}}$ , the squared pairwise Euclidean distance between two arbitrary noiseless receive vectors is given by:

$$d_{ij}^2 = \text{tr}(\mathbf{H}_{\mathbf{w}}\mathbf{P}(\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^\dagger \mathbf{P}^\dagger \mathbf{H}_{\mathbf{w}}^\dagger) \quad (141)$$

and its probability density function is given by:

$$p_{d_{ij}^2}(d_{ij}^2) = \frac{1}{(n_r - 1)!} \cdot \frac{1}{\left(\text{tr}\left((\mathbf{x}_i - \mathbf{x}_j)^\dagger \mathbf{P}^\dagger \mathbf{P}(\mathbf{x}_i - \mathbf{x}_j)\right)\right)^{n_r}} \cdot (d_{ij}^2)^{n_r - 1} \cdot e^{-\frac{d_{ij}^2}{\text{tr}\left((\mathbf{x}_i - \mathbf{x}_j)^\dagger \mathbf{P}^\dagger \mathbf{P}(\mathbf{x}_i - \mathbf{x}_j)\right)}}, \quad d_{ij}^2 \geq 0 \quad (142)$$

This, together with Theorem 2, leads to the high-snr characterization of the constrained capacity, which is a function of the precoder, given by:

$$\bar{I}(\text{snr}) = \log M - \epsilon'_{n_r}(\text{snr}) \cdot \frac{1}{\text{snr}^{n_r}} + \mathcal{O}\left(\frac{1}{\text{snr}^{n_r+1}}\right) \quad (143)$$

<sup>9</sup>We consider without any loss of generality the matrix  $\mathbf{P}$  to be square.

where

$$\begin{aligned}
& k'_{UB_{n_r}} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \left( \frac{1}{\text{tr} \left( (\mathbf{x}_i - \mathbf{x}_j)^\dagger \mathbf{P}^\dagger \mathbf{P} (\mathbf{x}_i - \mathbf{x}_j) \right)} \right)^{n_r} \\
& \leq \epsilon'_{n_r}(\text{snr}) \leq \\
& k'_{LB_{n_r}} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \left( \frac{1}{\text{tr} \left( (\mathbf{x}_i - \mathbf{x}_j)^\dagger \mathbf{P}^\dagger \mathbf{P} (\mathbf{x}_i - \mathbf{x}_j) \right)} \right)^{n_r}
\end{aligned} \tag{144}$$

and

$$k'_{UB_{n_r}} = \frac{1}{2M(M-1)} \cdot \frac{1}{\sqrt{\pi}} \cdot 4^n \cdot \frac{\Gamma(n_r + 3/2)}{\Gamma(n_r + 2)} \tag{145}$$

$$k'_{LB_{n_r}} = \frac{2}{M} \cdot \frac{1}{\sqrt{\pi}} \cdot 4^n \cdot \frac{\Gamma(n_r + 3/2)}{\Gamma(n_r + 2)} \tag{146}$$

Therefore, we pose the optimization problem, which is equivalent to the maximization of the lower bound (or the upper bound) to the constrained capacity of the canonical i.i.d. Rayleigh fading coherent channel in the asymptotic regime of high snr subject to a total power constraint, given by:

$$\min_{\mathbf{P}} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \left( \frac{1}{\text{tr} \left( (\mathbf{x}_i - \mathbf{x}_j)^\dagger \mathbf{P}^\dagger \mathbf{P} (\mathbf{x}_i - \mathbf{x}_j) \right)} \right)^{n_r} \tag{147}$$

with:

$$\text{tr}(\mathbf{P}\mathbf{P}^\dagger) \leq P \tag{148}$$

This optimization problem leads immediately to a simple precoder design procedure, that bypasses the need for time-consuming Monte Carlo simulations, based on numerical or analytical techniques. We illustrate the analysis by assuming that the individual precoder inputs conform to the same equiprobable symmetric constellation, e.g., some PSK or some QAM constellation.<sup>10</sup>

By introducing the change of variables  $\mathbf{Z} = \mathbf{P}^\dagger \mathbf{P}$ , we pose an optimization problem equivalent to the optimization problem in (147) and (148) as follows:

$$\min_{\mathbf{Z}} \sum_{\substack{i,j=1 \\ i \neq j}}^M \left( \frac{1}{\text{tr} \left( (\mathbf{x}_i - \mathbf{x}_j)^\dagger \mathbf{Z} (\mathbf{x}_i - \mathbf{x}_j) \right)} \right)^{n_r} \tag{149}$$

subject to:

$$\text{tr}(\mathbf{Z}) \leq P \tag{150}$$

$$\mathbf{Z} \succeq \mathbf{0} \tag{151}$$

<sup>10</sup>Note that this analysis only seems to be applicable to scenarios where the individual precoder inputs conform to the same equiprobable symmetric constellation, rather than scenarios where the individual precoder inputs conform to distinct constellations.

Note that this represents a convex optimization problem because: *i*) the objective function is convex in  $\mathbf{Z} \succeq \mathbf{0}$  ( $\mathbf{Z} \neq \mathbf{0}$ ); *ii*) the constraint set is convex. Define the Lagrangian of the optimization problem as follows:

$$\mathcal{L}(\mathbf{Z}, \mathbf{\Psi}, \lambda) = \sum_{\substack{i,j=1 \\ i \neq j}}^M \left( \frac{1}{\text{tr}((\mathbf{x}_i - \mathbf{x}_j) \mathbf{Z} (\mathbf{x}_i - \mathbf{x}_j)^\dagger)} \right)^{n_r} - \lambda \cdot (\mathbf{P} - \text{tr}(\mathbf{Z})) - \text{tr}(\mathbf{\Psi} \mathbf{Z}) \quad (152)$$

The Karush-Kuhn-Tucker conditions, which are both necessary and sufficient, state that the optimal solution  $\mathbf{Z}^*$  is such that:

$$\begin{aligned} \nabla_{\mathbf{Z}} \mathcal{L}(\mathbf{Z}, \mathbf{\Psi}, \lambda) |_{\mathbf{Z}=\mathbf{Z}^*} = \\ = -n_r \cdot \sum_{\substack{i,j=1 \\ i \neq j}}^M \left( \frac{1}{\text{tr}((\mathbf{x}_i - \mathbf{x}_j) \mathbf{Z}^* (\mathbf{x}_i - \mathbf{x}_j)^\dagger)} \right)^{n_r+1} \cdot (\mathbf{x}_i - \mathbf{x}_j) (\mathbf{x}_i - \mathbf{x}_j)^\dagger + \lambda \cdot \mathbf{I} - \mathbf{\Psi} = 0 \end{aligned} \quad (153)$$

$$\text{tr}(\mathbf{\Psi} \mathbf{Z}^*) = 0, \quad \mathbf{\Psi} \succeq \mathbf{0}, \quad \mathbf{Z}^* \succeq \mathbf{0} \quad (154)$$

$$\lambda \cdot (\mathbf{P} - \text{tr}(\mathbf{Z}^*)) = 0, \quad \lambda \geq 0 \quad (155)$$

It is possible to show that  $\mathbf{Z}^* = \frac{\mathbf{P}}{n_t} \cdot \mathbf{I}$  without any loss of generality, because  $\mathbf{Z}^* = \frac{\mathbf{P}}{n_t} \cdot \mathbf{I}$  together with

$$\lambda = n_r \cdot \sum_{\substack{i,j=1 \\ j \neq i}}^M \left( \frac{n_t/\mathbf{P}}{\text{tr}((\mathbf{x}_i - \mathbf{x}_j) (\mathbf{x}_i - \mathbf{x}_j)^\dagger)} \right)^{n_r+1} \cdot |\mathbf{x}_i(k) - \mathbf{x}_j(k)|^2 \quad (156)$$

which does not depend on  $k$  due to the symmetry conditions, and  $\mathbf{\Psi} = \mathbf{0}$  satisfy the Karush-Kuhn-Tucker conditions.

Note that  $\mathbf{x}(k)$  represents the  $k$ th element of the vector  $\mathbf{x}$ . It is only necessary to prove that:

$$n_r \cdot \sum_{\substack{i,j=1 \\ j \neq i}}^M \left( \frac{n_t/\mathbf{P}}{\text{tr}((\mathbf{x}_i - \mathbf{x}_j) (\mathbf{x}_i - \mathbf{x}_j)^\dagger)} \right)^{n_r+1} \cdot (\mathbf{x}_i - \mathbf{x}_j) (\mathbf{x}_i - \mathbf{x}_j)^\dagger = \lambda \cdot \mathbf{I} \quad (157)$$

Note that the element in the  $m$ -th row and  $n$ -th column of the matrix on the left hand side of (157) is given by:

$$\sum_{\substack{i,j=1 \\ j \neq i}}^M \left( \frac{n_t/\mathbf{P}}{\text{tr}((\mathbf{x}_i - \mathbf{x}_j) (\mathbf{x}_i - \mathbf{x}_j)^\dagger)} \right)^{n_r+1} \cdot (\mathbf{x}_i(m) - \mathbf{x}_j(m)) (\mathbf{x}_i(n) - \mathbf{x}_j(n))^\dagger \quad (158)$$

In view of the symmetry conditions, for fixed  $m$  and  $n$  it is possible to divide the set of pairs of indices  $\{(i, j) : i, j = 1, \dots, M, i \neq j\}$ , which contains  $M \cdot (M - 1)$  indices pairs, into  $\frac{1}{2} \cdot M \cdot (M - 1)$  sets of pairs of indices  $\{(i_1, j_1), (i_2, j_2)\}$ , which contain only two indices pairs, with the property that  $\mathbf{x}_{i_1}(m) = -\mathbf{x}_{i_2}(m)$ ,  $\mathbf{x}_{j_1}(m) = -\mathbf{x}_{j_2}(m)$ ,  $\mathbf{x}_{i_1}(k) = \mathbf{x}_{i_2}(k)$ ,  $\forall k \neq m$ , and  $\mathbf{x}_{j_1}(k) = \mathbf{x}_{j_2}(k)$ ,  $\forall k \neq m$ . Then, it follows immediately that:

$$\begin{aligned} & \left( \frac{n_t/\mathbf{P}}{\text{tr}((\mathbf{x}_{i_1} - \mathbf{x}_{j_1}) (\mathbf{x}_{i_1} - \mathbf{x}_{j_1})^\dagger)} \right)^{n_r+1} \cdot (\mathbf{x}_{i_1}(m) - \mathbf{x}_{j_1}(m)) (\mathbf{x}_{i_1}(n) - \mathbf{x}_{j_1}(n))^\dagger + \\ & \left( \frac{n_t/\mathbf{P}}{\text{tr}((\mathbf{x}_{i_2} - \mathbf{x}_{j_2}) (\mathbf{x}_{i_2} - \mathbf{x}_{j_2})^\dagger)} \right)^{n_r+1} \cdot (\mathbf{x}_{i_2}(m) - \mathbf{x}_{j_2}(m)) (\mathbf{x}_{i_2}(n) - \mathbf{x}_{j_2}(n))^\dagger = 0 \end{aligned} \quad (159)$$

and that

$$\sum_{\substack{i,j=1 \\ j \neq i}}^M \left( \frac{n_t/P}{\text{tr}((\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^\dagger)} \right)^{n_r+1} \cdot (\mathbf{x}_i(m) - \mathbf{x}_j(m))(\mathbf{x}_i(n) - \mathbf{x}_j(n))^\dagger = 0 \quad (160)$$

Note also that the  $n$ -th diagonal element of the matrix on the left hand side of (157) is given by:

$$\sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \left( \frac{n_t/P}{\text{tr}((\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^\dagger)} \right)^{n_r+1} \cdot (\mathbf{x}_i(n) - \mathbf{x}_j(n))(\mathbf{x}_i(n) - \mathbf{x}_j(n))^\dagger \quad (161)$$

In view of the symmetry conditions, it also follows immediately that this quantity is independent of  $n$ .

The fact that  $\mathbf{Z}^* = \frac{P}{n_t} \cdot \mathbf{I}$  without any loss of generality leads directly to the form of the precoder that maximizes the lower bound to the constrained capacity as follows:

$$\mathbf{P}^* = \frac{P}{n_t} \cdot \mathbf{Q} \mathbf{I} \quad (162)$$

where  $\mathbf{Q}$  is any unitary matrix. This result is very intuitive due to the symmetry of the system model.

Figure 18 shows that the constrained capacity achieved by the optimal precoder, obtained via direct optimization of the Monte-Carlo simulated constrained capacity, appears to be equal to the constrained capacity achieved by the high-snr isotropic precoder across the signal-to-noise ratio range. This suggests that the approach is sharp in the sense that it captures well the effect of the precoder on the constrained capacity of the canonical i.i.d. Rayleigh fading coherent channel.

2) *The Antenna-Correlated Rayleigh Fading Channel:* In this channel model, where  $\mathbf{H} = \Theta_{\mathbf{R}}^{\frac{1}{2}} \mathbf{H}_{\mathbf{w}} \Theta_{\mathbf{T}}^{\frac{1}{2}}$ , the squared pairwise Euclidean distance between two arbitrary noiseless receive vectors is given by:

$$d_{ij}^2 = \text{tr} \left( \Theta_{\mathbf{R}}^{\frac{1}{2}} \mathbf{H}_{\mathbf{w}} \Theta_{\mathbf{T}}^{\frac{1}{2}} \mathbf{P} (\mathbf{x}_i - \mathbf{x}_j) (\mathbf{x}_i - \mathbf{x}_j)^\dagger \mathbf{P}^\dagger \Theta_{\mathbf{T}}^{\frac{1}{2}} \mathbf{H}_{\mathbf{w}}^\dagger \Theta_{\mathbf{R}}^{\frac{1}{2}} \right) \quad (163)$$

and its probability density function is given by:

$$p_{d_{ij}^2}(d_{ij}^2) = \frac{1}{(n_r - 1)!} \cdot \frac{1}{\left( \text{tr}((\mathbf{x}_i - \mathbf{x}_j)^\dagger \mathbf{P}^\dagger \Theta_{\mathbf{T}} \mathbf{P} (\mathbf{x}_i - \mathbf{x}_j)) \right)^{n_r}} \cdot (d_{ij}^2)^{n_r-1} \cdot e^{-\frac{d_{ij}^2}{\text{tr}((\mathbf{x}_i - \mathbf{x}_j)^\dagger \mathbf{P}^\dagger \Theta_{\mathbf{T}} \mathbf{P} (\mathbf{x}_i - \mathbf{x}_j))}}, \quad d_{ij}^2 \geq 0 \quad (164)$$

when the eigenvalues of the receive correlation matrix,  $\lambda_{R_1}, \lambda_{R_2}, \dots, \lambda_{R_{n_r}}$ , are all equal (necessarily to one) or

$$p_{d_{ij}^2}(d_{ij}^2) = \sum_{k=1}^{n_r} \frac{1}{\lambda_{R_k} \prod_{\substack{k'=1 \\ k' \neq k}}^{n_r} \left( 1 - \frac{\lambda_{R_{k'}}}{\lambda_{R_k}} \right) \cdot \text{tr}((\mathbf{x}_i - \mathbf{x}_j)^\dagger \mathbf{P}^\dagger \Theta_{\mathbf{T}} \mathbf{P} (\mathbf{x}_i - \mathbf{x}_j))} \cdot e^{-\frac{d_{ij}^2}{\lambda_{R_k} \cdot \text{tr}((\mathbf{x}_i - \mathbf{x}_j)^\dagger \mathbf{P}^\dagger \Theta_{\mathbf{T}} \mathbf{P} (\mathbf{x}_i - \mathbf{x}_j))}}, \quad d_{ij}^2 \geq 0 \quad (165)$$

when the eigenvalues of the receive correlation matrix,  $\lambda_{R_1}, \lambda_{R_2}, \dots, \lambda_{R_{n_r}}$ , are all distinct<sup>11</sup>.

<sup>11</sup>We once again concentrate on non-degenerate scenarios where  $\lambda_{T_{ij}} > 0, \forall i \neq j$ , and  $\lambda_{R_k} > 0, \forall k$ , and on scenarios where the eigenvalues  $\lambda_{R_k}, k = 1, \dots, n_r$ , are either all distinct or all equal (necessarily to one).

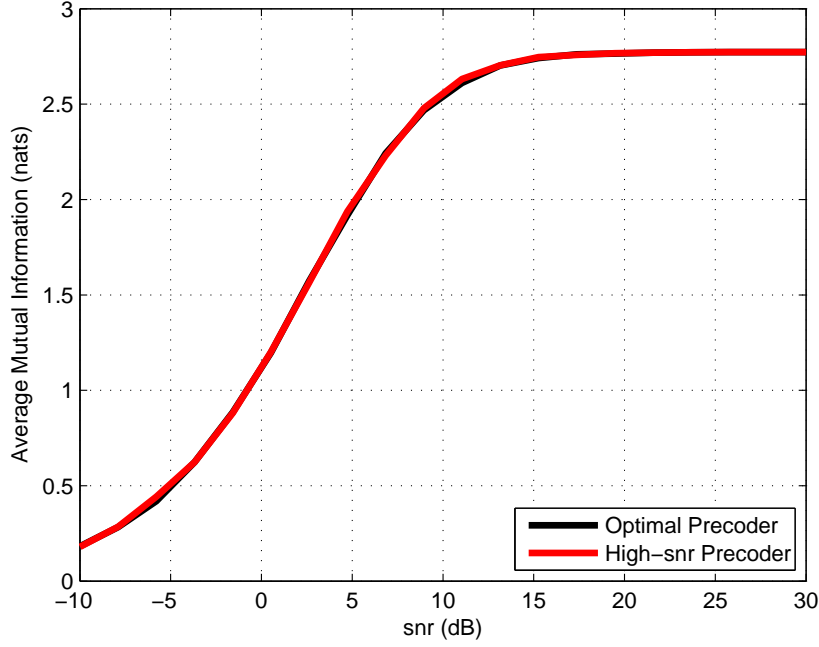


Fig. 18. Average mutual information for a  $2 \times 2$  canonical i.i.d. Rayleigh fading coherent channel driven by QPSK inputs.

This, together with Theorem 2, also leads to the high-snr characterization of the constrained capacity, which is a function of the precoder, given by:

$$\bar{I}(\text{snr}) = \log M - \epsilon'_{n_r}(\text{snr}) \cdot \frac{1}{\text{snr}^{n_r}} + \mathcal{O}\left(\frac{1}{\text{snr}^{n_r+1}}\right) \quad (166)$$

where

$$\begin{aligned} & k'_{UB_{n_r}} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ i \neq j}}^M \left( \frac{1}{\text{tr} \left( (\mathbf{x}_i - \mathbf{x}_j)^\dagger \mathbf{P}^\dagger \boldsymbol{\Theta}_T \mathbf{P} (\mathbf{x}_i - \mathbf{x}_j) \right)} \right)^{n_r} \cdot \frac{1}{\det(\boldsymbol{\Theta}_R)} \\ & \leq \epsilon'_{n_r}(\text{snr}) \leq \\ & k'_{LB_{n_r}} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ i \neq j}}^M \left( \frac{1}{\text{tr} \left( (\mathbf{x}_i - \mathbf{x}_j)^\dagger \mathbf{P}^\dagger \boldsymbol{\Theta}_T \mathbf{P} (\mathbf{x}_i - \mathbf{x}_j) \right)} \right)^{n_r} \cdot \frac{1}{\det(\boldsymbol{\Theta}_R)} \end{aligned} \quad (167)$$

and

$$k'_{UB_{n_r}} = \frac{1}{2M(M-1)} \cdot \frac{1}{\sqrt{\pi}} \cdot 4^n \cdot \frac{\Gamma(n_r + 3/2)}{\Gamma(n_r + 2)} \quad (168)$$

$$k'_{LB_{n_r}} = \frac{2}{M} \cdot \frac{1}{\sqrt{\pi}} \cdot 4^n \cdot \frac{\Gamma(n_r + 3/2)}{\Gamma(n_r + 2)} \quad (169)$$

Therefore, we pose the optimization problem, which is equivalent to the maximization of the lower bound (or the upper bound) to the constrained capacity of the antenna-correlated Rayleigh fading coherent channel in the

asymptotic regime of high snr subject to a total power constraint, given by:

$$\min_{\mathbf{P}} \sum_{i=1}^M \sum_{\substack{j=1 \\ i \neq j}}^M \left( \frac{1}{\text{tr} \left( (\mathbf{x}_i - \mathbf{x}_j)^\dagger \mathbf{P}^\dagger \boldsymbol{\Theta}_T \mathbf{P} (\mathbf{x}_i - \mathbf{x}_j) \right)} \right)^{n_r} \cdot \frac{1}{\det(\boldsymbol{\Theta}_R)} \quad (170)$$

with:

$$\text{tr}(\mathbf{P}\mathbf{P}^\dagger) \leq P \quad (171)$$

This optimization problem, akin to the previous one, also leads immediately to a simple precoder design procedure based on numerical techniques. Most notably, it is possible to prove from (170) and (171), using the techniques in [35], that the matrix containing the left singular vectors of the precoder matrix corresponds to the matrix containing the right singular vectors of the transmit-antennas correlation matrix, i.e., the optimal precoder diagonalizes the transmit-antenna correlation matrix. It is also possible to prove from (170) and (171) that, upon setting the matrix containing the left singular vectors of the precoder matrix to be equal to the matrix containing the right singular vectors of the transmit-antenna correlation matrix, the optimization problem becomes concave in the precoder squared singular values. These facts have also been recently established by different means in [36].

Figure 19 shows that the constrained capacity achieved by the optimal precoder design, obtained via direct optimization of the Monte-Carlo simulated constrained capacity, tends with the increase of the signal-to-noise ratio to the constrained capacity achieved by the high-snr precoder design, obtained via the optimization problem in (170) and (171). This suggests once again that the design approach is sharp in the sense that it also captures well the effect of the precoder on the constrained capacity of the antenna-correlated Rayleigh fading coherent channel.

### C. Space-Time Coding in the Multiple-Antenna Canonical i.i.d. Rayleigh Fading Channel

We now consider a more general multiple-antenna fading channel, which encompasses communication over  $t$  symbol intervals, given by:

$$\mathbf{Y} = \sqrt{\text{snr}} \cdot \mathbf{H}_w \mathbf{X} + \mathbf{N} \quad (172)$$

where  $\mathbf{Y} \in \mathbb{C}^{n_r \times t}$  denotes the receive codeword matrix,  $\mathbf{X} \in \mathbb{C}^{n_t \times t}$  denotes the transmit codeword matrix,  $\mathbf{H}_w \in \mathbb{C}^{n_r \times n_t}$  is the canonical i.i.d. complex Gaussian random channel fading matrix, and  $\mathbf{N} \in \mathbb{C}^{n_r \times t}$  is the noise matrix consisting of independent circularly symmetric complex Gaussian random variables with zero-mean and unit-variance. We assume that the channel matrix is constant for the duration of a codeword, changing from codeword to codeword in a stationary and ergodic manner. We also assume that the receiver knows the exact channel matrix realization but the transmitter knows only the channel matrix distribution. The space-time codeword matrices conform to an equiprobable multi-dimensional constellation with cardinality  $M$ , i.e.,  $\mathbf{X} \in \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_M\}$  and  $\Pr(\mathbf{X}_1) = \Pr(\mathbf{X}_2) = \dots = \Pr(\mathbf{X}_M) = \frac{1}{M}$ .

By capitalizing on Theorem 2, the high-snr expansion of the constrained capacity in nats per channel use, where a channel use encompasses  $t$  symbol intervals, obeys:

$$\bar{I}(\text{snr}) = \log M - \epsilon'_d(\text{snr}) \cdot \frac{1}{\text{snr}^d} + \mathcal{O}\left(\frac{1}{\text{snr}^{d+1}}\right) \quad (173)$$

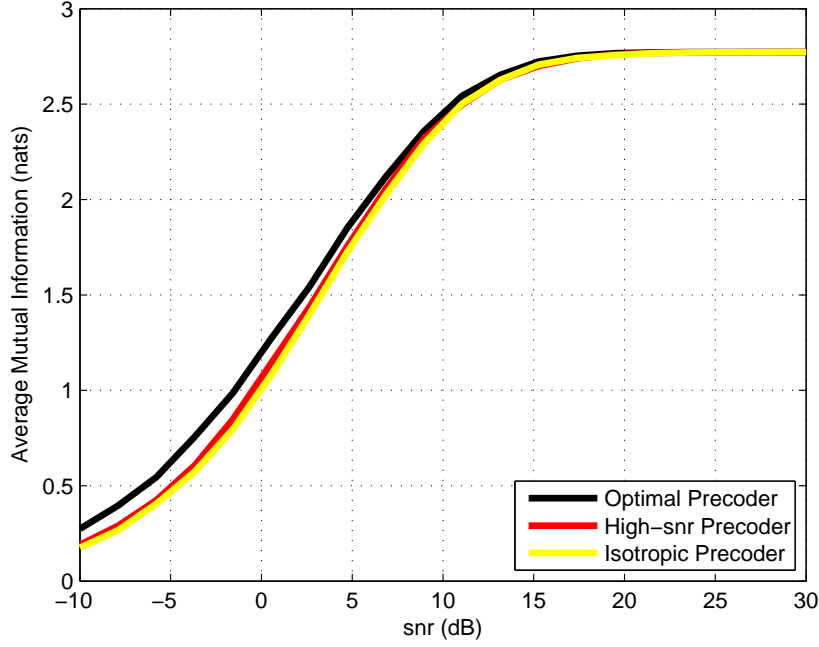


Fig. 19. Average mutual information for a  $2 \times 2$  antenna-correlated Rayleigh fading coherent channel driven by QPSK inputs ( $\Phi_{\mathbf{T}} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$  and  $\Phi_{\mathbf{R}} = \begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix}$ ).

where

$$d = 1 + \min \left\{ n \in \mathbb{N}_0 : \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(n)}(0) \neq 0 \right\} \quad (174)$$

and

$$k'_{\text{UB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) \leq \epsilon'_d(\text{snr}) \leq k'_{\text{LB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) \quad (175)$$

$$k'_{\text{LB}_d} = \frac{2}{M} \cdot \frac{4^d}{\sqrt{\pi}} \cdot \frac{\Gamma(d+3/2)}{\Gamma(d+2)} \quad (176)$$

$$k'_{\text{UB}_d} = \frac{1}{2M(M-1)} \cdot \frac{4^d}{\sqrt{\pi}} \cdot \frac{\Gamma(d+3/2)}{\Gamma(d+2)} \quad (177)$$

and  $p_{d_{ij}^2}(\cdot)$  and  $p_{d_{ij}^2}^{(n)}(\cdot)$  denote the probability density function and the higher-order derivatives of the probability density function of the squared Euclidean distance between pairs of arbitrary noiseless receive codeword matrices.

It is evident that the exact value of  $d$  and the bounds to the value of  $\epsilon'_d(\text{snr})$  are connected to the properties of the set of space-time codewords. The objective is to unveil the structure of the set of space-time codewords that lead to: *i*) the maximization of the exact value of  $d$ ; and *ii*) the minimization of the upper bound (as well as the

lower bound) to the value of  $\epsilon'_d(\text{snr})$ . This leads to the maximization of the lower bound to the constrained capacity in the asymptotic regime of high-snr regime.

The following Theorem summarizes the space-time code design criteria for the canonical i.i.d. multiple-antenna Rayleigh fading coherent channel, with two transmit antennas and one or two receive antennas. We denote  $\Delta_{ij} = (\mathbf{X}_i - \mathbf{X}_j)(\mathbf{X}_i - \mathbf{X}_j)^\dagger$ . We also denote the eigenvalue decomposition of the matrix  $\Delta_{ij}$  by  $\Delta_{ij} = \mathbf{W}\Lambda_{ij}\mathbf{W}^\dagger$ , where  $\mathbf{W}$  is a unitary matrix,  $\Lambda_{ij} = \text{diag}(\lambda_{ij}(1), \dots, \lambda_{ij}(r_{ij}), 0, \dots, 0)$  is a diagonal matrix,  $\lambda_{ij}(1) \geq \dots \geq \lambda_{ij}(r_{ij}) > 0$  are the non-zero eigenvalues of  $\Delta_{ij}$ , and  $r_{ij} = \text{rank}(\Delta_{ij}) \leq \min(n_t, t)$  is the rank of  $\Delta_{ij}$ . We let  $R_{ij} \leq r_{ij}$  represent the number of distinct eigenvalues in  $\lambda_{ij}(1), \dots, \lambda_{ij}(r_{ij})$ . We also let  $\lambda'_{ij}(1), \dots, \lambda'_{ij}(R_{ij})$  and  $r'_{ij}(1), \dots, r'_{ij}(R_{ij})$  represent the distinct eigenvalues and their multiplicities, respectively, where  $r'_{ij}(1) + \dots + r'_{ij}(R_{ij}) = r_{ij}$ .

*Theorem 6:* Consider the canonical i.i.d. multiple-antenna Rayleigh fading coherent channel model in (172) with  $n_t = 2$  and  $n_r = 1$  or  $n_r = 2$ , where  $\mathbf{X} \in \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_M\}$  and  $\Pr(\mathbf{X} = \mathbf{X}_1) = \Pr(\mathbf{X} = \mathbf{X}_2) = \dots = \Pr(\mathbf{X} = \mathbf{X}_M) = \frac{1}{M}$ . Then,

- The value of  $d$  is maximized by the set of space-time codewords that maximize:

$$r_{\min} = \min_{i \neq j} r_{ij} \quad (178)$$

- The value of the bounds to  $\epsilon'_d(\text{snr})$  are maximized by the set of space-time codewords that maximize:

$$\sum_{\{i,j\} \in \Omega} \prod_{r=1}^{r_{ij}} \left( \frac{1}{\lambda_{ij}(r)} \right)^{n_r} \quad (179)$$

where  $\Omega = \{\{i, j\} \in \{1, \dots, M\} \times \{1, \dots, M\} : r_{ij} = r_{\min}, i \neq j\}$ .

*Proof:* Let us consider the squared Euclidean distance between a pair of arbitrary (noiseless) receive codeword matrices associated with the pair of transmit codeword matrices  $\mathbf{X}_i$  and  $\mathbf{X}_j$  given by:

$$d_{ij}^2 = \text{tr}(\mathbf{H}_w \Delta_{ij} \mathbf{H}_w^\dagger) \quad (180)$$

It is straightforward to show that the distribution of

$$\text{tr}(\mathbf{H}_w \Delta_{ij} \mathbf{H}_w^\dagger) = \text{tr}(\mathbf{H}_w \mathbf{W} \Lambda_{ij} \mathbf{W}^\dagger \mathbf{H}_w^\dagger) \quad (181)$$

is equal to the distribution of

$$\text{tr}(\mathbf{H}_w \Lambda_{ij} \mathbf{H}_w^\dagger) = \sum_{m=1}^{n_r} \sum_{n=1}^{r_{ij}} \lambda_{ij}(n) |\xi_{mn}|^2 \quad (182)$$

where  $\xi_{mn}$  are independent circularly symmetric complex Gaussian random variables with zero-mean and unit-variance. Its probability density function is given by [37], [38]:

$$p_{d_{ij}^2}(d_{ij}^2) = \frac{1}{\prod_{r=1}^{R_{ij}} \lambda'_{ij}(r)^{n_r r'_{ij}(r)}} \cdot \sum_{r=1}^{R_{ij}} \sum_{r'=1}^{n_r r'_{ij}(r)} \frac{\Psi_{r,r'}(n_r r'_{ij}(1), \dots, n_r r'_{ij}(R_{ij}))}{(n_r r'_{ij}(r) - r')!} \cdot (-d_{ij}^2)^{n_r r'_{ij}(r) - r'} \cdot e^{-\frac{d_{ij}^2}{\lambda'_{ij}(r)}}, \quad d_{ij}^2 \geq 0 \quad (183)$$



where

$$\begin{aligned} \Psi_{r,r'}(n_r r'_{ij}(1), \dots, n_r r'_{ij}(R_{ij})) &= (-1)^{n_r r'_{ij}(r)-1} \sum_{k_1, \dots, k_{R_{ij}} \in \Omega_{r,r'}} \sum_{\substack{k=1 \\ k \neq r}}^{R_{ij}} \binom{k_k + n_r r'_{ij}(k) - 1}{k_k} \\ &\quad \cdot \left( \frac{1}{\lambda'_{ij}(k)} - \frac{1}{\lambda'_{ij}(r)} \right)^{-(n_r r'_{ij}(k) - k_k)} \end{aligned} \quad (184)$$

with

$$\Omega_{r,r'} = \{k_1, \dots, k_{R_{ij}} \in \mathbb{Z}^{R_{ij}} : k_1 + \dots + k_{R_{ij}} = r' - 1, k_r = 0, k_k \geq 0 \ \forall \ k\} \quad (185)$$

Let us consider the scenario where  $n_t = 2$  and  $n_r = 1$ . It is possible to show from (183), (184) and (185) that if  $\lambda_{ij}(1) > 0$  and  $\lambda_{ij}(2) > 0$ , which implies that  $r_{ij} = 2$ , then

$$p_{d_{ij}^2}(0) = 0 \quad (186)$$

and

$$p_{d_{ij}^2}^{(1)}(0) = \frac{1}{\lambda_{ij}(1)\lambda_{ij}(2)} \neq 0 \quad (187)$$

whereas if  $\lambda_{ij}(1) > 0$  and  $\lambda_{ij}(2) = 0$ , which implies that  $r_{ij} = 1$ , then

$$p_{d_{ij}^2}(0) = \frac{1}{\lambda_{ij}(1)} \neq 0 \quad (188)$$

Let us now consider the scenario where  $n_t = 2$  and  $n_r = 2$ . It is also possible to show from (183), (184) and (185) that if  $\lambda_{ij}(1) > 0$  and  $\lambda_{ij}(2) > 0$  then

$$p_{d_{ij}^2}^{(n)}(0) = 0, \quad n = 0, 1, 2 \quad (189)$$

and

$$p_{d_{ij}^2}^{(3)}(0) = \frac{1}{\lambda_{ij}^2(1)\lambda_{ij}^2(2)} \neq 0 \quad (190)$$

where as if  $\lambda_{ij}(1) > 0$  and  $\lambda_{ij}(2) = 0$  then

$$p_{d_{ij}^2}(0) = 0 \quad (191)$$

and

$$p_{d_{ij}^2}^{(1)}(0) = \frac{1}{\lambda_{ij}^2(1)} \neq 0 \quad (192)$$

Therefore, the Theorem follows immediately from the high-snr characterization in (173), (174), (175), (176) and (177). ■

In principle, it is also possible to generalize the result for canonical i.i.d. multiple-antenna Rayleigh fading coherent channels with an arbitrary number of transmit and receive antennas. However, this requires considerable algebraic manipulation due to the form of the probability density function in (183), (184) and (185).

Note that if  $n_t \geq t$ , then the maximum possible value for  $r_{ij}$  is  $t$ ; in contrast, if  $n_t \leq t$ , then the maximum possible value of  $r_{ij}$  is  $n_t$ . Note also that if  $r_{ij} = n_t, \forall i \neq j$ , then the maximization of:

$$\sum_{\{i,j\} \in \Omega} \prod_{r=1}^{r_{ij}} \left( \frac{1}{\lambda_{ij}(r)} \right)^{n_r} \quad (193)$$

corresponds to the maximization of:

$$\sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \left( \frac{1}{\det(\Delta_{ij})} \right)^{n_r} \quad (194)$$

We conclude that the space-time code design criteria embodied in Theorem 6 are akin to the conventional rank and determinant design criteria in [39], which have been derived from the pairwise error probability view point. Since the average error probability, the average MMSE and the average mutual information exhibit identical high-snr behavior, the design criteria put forth in Theorem 6, in addition to maximizing the lower bound to the average mutual information, also minimize the upper bounds to the average MMSE and the average error probability.

## VII. CONCLUSIONS

By drawing on the I-MMSE identity and counterparts, together with key results in asymptotic analysis and expansions, it has been possible to put forth a high-snr characterization of the constrained capacity of multiple-antenna fading coherent channels driven by arbitrary equiprobable discrete inputs. This characterization has enabled the study of the effect of various system models, parameters, and elements on the constrained capacity of multiple-antenna fading coherent channels. Key contributions include:

- In the antenna-uncorrelated Rayleigh fading coherent channel, we have analyzed the impact on the high-snr behavior of the constrained capacity of the number of transmit antennas, the number of receive antennas and the characteristics of the multi-dimensional constellation. It has been observed that in the regime of high-snr the constrained capacity increases with the increase in the number of transmit antennas and the number of receive antennas, as well as with the use of multi-dimensional constellations with better sphere-packing properties. In particular, it has been observed that in the regime of high-snr the number of receive antennas control the rate at which the constrained capacity value tends to its infinite-snr value whereas the number of transmit antennas control the constrained capacity power offset, by enabling the construction of multi-dimensional constellations with better sphere-packing properties. It has also been observed that the geometry of the multi-dimensional constellation, most notably, its sphere-packing properties, also control the constrained capacity power offset. It has also been emphasized that in the regime of high-snr the role of the transmit and receive antennas on the behavior of the constrained capacity of canonical i.i.d. Rayleigh fading channels is very different from their role on the behavior of the capacity of the canonical i.i.d. Rayleigh fading coherent channel, which is achieved by Gaussian inputs.
- In the antenna-correlated Rayleigh fading channel, we have analyzed the impact on the high-snr behavior of the constrained capacity of the number of transmit antennas, the number of receive antennas, transmit and receive antenna correlation and the characteristics of the multi-dimensional constellation. It has been observed

that in the regime of high-snr the constrained capacity also increases with the increase in the number of transmit antennas and the number of receive antennas, as well as with the use of multi-dimensional constellations with better sphere-packing properties in the coordinate system induced by the transmit correlation matrix. It has been found that the presence of transmit and receive antenna correlation has a negative impact on the constrained capacity in the regime of high-snr. It has also been found that degenerate conditions, i.e., perfectly correlated paths, have an impact on the constrained capacity infinite-snr value as well as the rate at which the constrained capacity value tends to its infinite-snr value.

- In Ricean fading channels line-of-sight components also have profound implications on the behavior of the constrained capacity in the regime of high snr. Of particular interest, it has been observed that, whereas in a single-transmit single-receive antenna Ricean fading coherent channel the constrained capacity increases with the increase in the K-factor in the regime of high-snr, in a multiple-transmit multiple-receive antenna Ricean fading channel the constrained capacity may decrease with the increase of the K-factor in the regime of high-snr. This result, which does not occur for Gaussian inputs, is due to the nature of the interaction of the multi-dimensional constellation with the channel model.
- The high-snr characterization of the constrained capacity has also enabled the design of elements for multiple-antenna fading coherent channel models. We have considered power allocation in a bank of parallel independent fading coherent channels, precoding for multiple-antenna Rayleigh fading coherent channels and space-time coding for multiple-antenna Rayleigh fading coherent channels, showing that the expansions lead to very sharp designs.

Of particular relevance, the construction of the high-snr constrained capacity characterization has also disclosed intimate connections between the high-snr asymptotic behavior of the average minimum mean-squared error, the average mutual information and the average error probability in multiple-antenna fading coherent channels driven by arbitrary equiprobable discrete inputs. These connections, which suggest that designs that minimize (bounds to) the error probability also minimize (bounds to) the minimum mean-squared error and maximize (bounds to) the average mutual information, provide a unification of the behavior of key performance metrics in the asymptotic regime of high snr. Overall, this high-snr analysis complements the low-snr insight available in [4], [5].

## APPENDIX A

### PROOF OF LEMMA 1

We determine the lower bound to the MMSE by using a genie that supplies the receiver with the true input vector and any of the other input vectors with equal probability. The genie based estimate is given by:

$$\hat{\mathbf{x}}_{genie}(\mathbf{y}, \{\mathbf{x}_i, \mathbf{x}_j\}) = \frac{\mathbf{x}_i e^{-\|\mathbf{y} - \sqrt{\text{snr}} \mathbf{H} \mathbf{x}_i\|^2} + \mathbf{x}_j e^{-\|\mathbf{y} - \sqrt{\text{snr}} \mathbf{H} \mathbf{x}_j\|^2}}{e^{-\|\mathbf{y} - \sqrt{\text{snr}} \mathbf{H} \mathbf{x}_i\|^2} + e^{-\|\mathbf{y} - \sqrt{\text{snr}} \mathbf{H} \mathbf{x}_j\|^2}} \quad (195)$$

where  $\mathbf{x}_i$  is the true input vector and  $\mathbf{x}_j$  is the other input vector. It follows that the MMSE can be lower bounded as follows:

$$\text{mmse}(\text{snr}; \mathbf{H}) = \mathbb{E} \left\{ \|\mathbf{H}\mathbf{x} - \mathbf{H}\mathbb{E}\{\mathbf{x}|\mathbf{y}\}\|^2 | \mathbf{H} \right\} \quad (196)$$

$$= \frac{1}{M} \sum_{i=1}^M \mathbb{E} \left\{ \|\mathbf{H}\mathbf{x} - \mathbf{H}\mathbb{E}\{\mathbf{x}|\mathbf{y}\}\|^2 | \mathbf{x} = \mathbf{x}_i, \mathbf{H} \right\} \quad (197)$$

$$\geq \frac{1}{M(M-1)} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \mathbb{E} \left\{ \|\mathbf{H}\mathbf{x} - \mathbf{H}\hat{\mathbf{x}}_{\text{genie}}(\mathbf{y}, \{\mathbf{x}_i, \mathbf{x}_j\})\|^2 | \mathbf{x} = \mathbf{x}_i, \mathbf{H} \right\} \quad (198)$$

$$= \frac{1}{M(M-1)} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \frac{\|\mathbf{H}\mathbf{x}_i - \mathbf{H}\mathbf{x}_j\|^2}{4} \cdot \text{mmse}_{\text{BPSK}} \left( \frac{\|\mathbf{H}\mathbf{x}_i - \mathbf{H}\mathbf{x}_j\|^2 \text{snr}}{4} \right) \quad (199)$$

$$= \frac{1}{M(M-1)} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \frac{d_{ij}^2(\mathbf{H})}{4} \cdot \text{mmse}_{\text{BPSK}} \left( \frac{d_{ij}^2(\mathbf{H}) \text{snr}}{4} \right) \quad (200)$$

$$\geq \frac{1}{M(M-1)} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \frac{d_{ij}^2(\mathbf{H})}{4} \cdot \frac{1}{2} \cdot \text{erfc} \left( \sqrt{\frac{d_{ij}^2(\mathbf{H}) \text{snr}}{4}} \right) \triangleq \text{mmse}_{\text{LB}}(\text{snr}; \mathbf{H}) \quad (201)$$

The first inequality in (198) is due to the genie based estimator and the second inequality in (201) is due to the lower bound to the MMSE of BPSK:

$$\text{mmse}_{\text{BPSK}}(\rho) = 1 - \int_{-\infty}^{\infty} \tanh(2\sqrt{\rho}\xi) \frac{e^{-(\xi-\sqrt{\rho})^2}}{\sqrt{\pi}} d\xi \geq 1 - \int_0^{\infty} \frac{e^{-(\xi-\sqrt{\rho})^2}}{\sqrt{\pi}} d\xi = 1 - \int_{-\sqrt{\rho}}^{\infty} \frac{e^{-\xi^2}}{\sqrt{\pi}} d\xi = \frac{1}{2} \cdot \text{erfc}(\sqrt{\rho}) \quad (202)$$

We determine the upper bound to the MMSE by using a (sub-optimal) Euclidean distance based estimator rather than the optimal conditional mean estimator. The Euclidean distance based estimate is given by:

$$\hat{\mathbf{x}}_{\text{euc}}(\mathbf{y}) = \underset{\mathbf{x}}{\text{argmin}} \|\mathbf{y} - \sqrt{\text{snr}} \cdot \mathbf{H}\mathbf{x}\|^2, \quad (203)$$

It follows that the MMSE can be upper bounded as follows:

$$\text{mmse}(\text{snr}; \mathbf{H}) = \mathbb{E} \left\{ \|\mathbf{H}\mathbf{x} - \mathbf{H}\mathbb{E}\{\mathbf{x}|\mathbf{y}\}\|^2 | \mathbf{H} \right\} \quad (204)$$

$$= \frac{1}{M} \sum_{i=1}^M \mathbb{E} \left\{ \|\mathbf{H}\mathbf{x} - \mathbf{H}\mathbb{E}\{\mathbf{x}|\mathbf{y}\}\|^2 | \mathbf{x} = \mathbf{x}_i, \mathbf{H} \right\} \quad (205)$$

$$\leq \frac{1}{M} \sum_{i=1}^M \mathbb{E} \left\{ \|\mathbf{H}\mathbf{x} - \mathbf{H}\hat{\mathbf{x}}_{\text{euc}}(\mathbf{y})\|^2 | \mathbf{x} = \mathbf{x}_i, \mathbf{H} \right\} \quad (206)$$

$$= \frac{1}{M} \sum_{i=1}^M \sum_{j=1}^M \mathbb{E} \left\{ \|\mathbf{H}\mathbf{x} - \mathbf{H}\hat{\mathbf{x}}_{\text{euc}}(\mathbf{y})\|^2 | \mathbf{x} = \mathbf{x}_i, \mathbf{y} \in \mathcal{V}_j, \mathbf{H} \right\} \cdot \Pr \{ \mathbf{y} \in \mathcal{V}_j | \mathbf{x} = \mathbf{x}_i, \mathbf{H} \} \quad (207)$$

$$\leq \frac{1}{M} \sum_{i=1}^M \sum_{j=1}^M d_{ij}^2(\mathbf{H}) \cdot \frac{1}{2} \cdot \text{erfc} \left( \sqrt{\frac{d_{ij}^2(\mathbf{H}) \text{snr}}{4}} \right) \triangleq \text{mmse}_{\text{UB}}(\text{snr}; \mathbf{H}) \quad (208)$$

where  $\mathcal{V}_l$  is the Voronoi region associated with  $\mathbf{H}\mathbf{x}_l$ . The first inequality in (206) is due to the (sub-optimal) Euclidean distance based estimator and the second inequality in (208) is due to

$$\Pr \{ \mathbf{y} \in \mathcal{V}_j | \mathbf{x} = \mathbf{x}_i, \mathbf{H} \} \leq \frac{1}{2} \cdot \text{erfc} \left( \sqrt{\frac{d_{ij}^2(\mathbf{H}) \text{snr}}{4}} \right) \quad (209)$$

and  $\mathbb{E} \left\{ \|\mathbf{H}\mathbf{x} - \mathbf{H}\hat{\mathbf{x}}_{\text{euc}}(\mathbf{y})\|^2 | \mathbf{x} = \mathbf{x}_i, \mathbf{y} \in \mathcal{V}_j, \mathbf{H} \right\} = d_{ij}^2(\mathbf{H})$ .

## APPENDIX B

### PROOF OF LEMMA 2

We determine the lower bound to the mutual information by using (15) in (16) as follows:

$$I(\text{snr}; \mathbf{H}) \geq \log M - \int_{\text{snr}}^{\infty} \text{mmse}_{\text{UB}}(\xi; \mathbf{H}) d\xi \quad (210)$$

$$= \log M - \frac{1}{M} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \int_{\text{snr}}^{\infty} d_{ij}^2(\mathbf{H}) \cdot \frac{1}{2} \cdot \text{erfc} \left( \sqrt{\frac{d_{ij}^2(\mathbf{H}) \xi}{4}} \right) d\xi \quad (211)$$

$$\geq \log M - \frac{1}{M} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \int_{\text{snr}}^{\infty} d_{ij}^2(\mathbf{H}) \cdot \frac{1}{2} \cdot e^{-\frac{d_{ij}^2(\mathbf{H}) \xi}{4}} d\xi \quad (212)$$

$$= \log M - \frac{1}{M} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M 2 \cdot e^{-\frac{d_{ij}^2(\mathbf{H}) \text{snr}}{4}} \triangleq I_{\text{LB}}(\text{snr}; \mathbf{H}) \quad (213)$$

The inequality in (212) follows from the upper bound to the complementary error function  $\text{erfc}(x) \leq e^{-x^2}$ .

We determine the upper bound to the mutual information by using (14) in (16) as follows:

$$I(\text{snr}; \mathbf{H}) \leq \log M - \int_{\text{snr}}^{\infty} \text{mmse}_{\text{LB}}(\xi; \mathbf{H}) d\xi \quad (214)$$

$$= \log M - \frac{1}{M(M-1)} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \int_{\text{snr}}^{\infty} \frac{d_{ij}^2(\mathbf{H})}{4} \cdot \frac{1}{2} \cdot \text{erfc} \left( \sqrt{\frac{d_{ij}^2(\mathbf{H}) \xi}{4}} \right) d\xi \quad (215)$$

$$= \log M + \frac{1}{M(M-1)} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \frac{d_{ij}^2(\mathbf{H}) \text{snr}}{4} \cdot \frac{1}{2} \cdot \text{erfc} \left( \sqrt{\frac{d_{ij}^2(\mathbf{H}) \text{snr}}{4}} \right) \quad (216)$$

$$- \frac{1}{M(M-1)} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \sqrt{\frac{d_{ij}^2(\mathbf{H}) \text{snr}}{4}} \cdot \frac{1}{2} \cdot \frac{e^{-\frac{d_{ij}^2(\mathbf{H}) \text{snr}}{4}}}{\sqrt{\pi}} \quad (217)$$

$$- \frac{1}{M(M-1)} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \frac{1}{4} \cdot \text{erfc} \left( \sqrt{\frac{d_{ij}^2(\mathbf{H}) \text{snr}}{4}} \right) \quad (218)$$

$$\leq \log M - \frac{1}{M(M-1)} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \frac{1}{4} \cdot \text{erfc} \left( \sqrt{\frac{d_{ij}^2(\mathbf{H}) \text{snr}}{4}} \right) \triangleq I_{\text{UB}}(\text{snr}; \mathbf{H}) \quad (219)$$

The inequality in (219) follows from the upper bound to the complementary error function  $\text{erfc}(x) \leq e^{-x^2}/\sqrt{\pi \cdot x^2}$ .

APPENDIX C  
PROOF OF LEMMA 3

The lower and upper bounds to the average value of the MMSE are obtained from the lower and upper bounds to the MMSE, respectively, by averaging over the fading statistics as follows:

$$\begin{aligned}\overline{\text{mmse}}_{\text{LB}}(\text{snr}) &= \mathbb{E}_{\mathbf{H}} \{ \text{mmse}_{\text{LB}}(\text{snr}; \mathbf{H}) \} \\ &= \frac{1}{\text{snr}} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \int_0^\infty \frac{1}{8M(M-1)} \cdot d_{ij}^2 \text{snr} \cdot \text{erfc} \left( \sqrt{\frac{d_{ij}^2 \text{snr}}{4}} \right) \cdot p_{d_{ij}^2}(d_{ij}^2) dd_{ij}^2\end{aligned}\quad (220)$$

$$\begin{aligned}\overline{\text{mmse}}_{\text{UB}}(\text{snr}) &= \mathbb{E}_{\mathbf{H}} \{ \text{mmse}_{\text{UB}}(\text{snr}; \mathbf{H}) \} \\ &= \frac{1}{\text{snr}} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \int_0^\infty \frac{1}{2M} \cdot d_{ij}^2 \text{snr} \cdot \text{erfc} \left( \sqrt{\frac{d_{ij}^2 \text{snr}}{4}} \right) \cdot p_{d_{ij}^2}(d_{ij}^2) dd_{ij}^2\end{aligned}\quad (221)$$

We establish the asymptotic expansions as  $\text{snr} \rightarrow \infty$  of the integrals that compose the lower bound and the upper bound to the average value of the MMSE in (220) and (221), respectively, by using [29, Theorem 3.2]. This Theorem requires that:

$$\left| \int_{\lambda t}^\infty dt_{n-1} \int_{t_{n-1}}^\infty dt_{n-2} \cdots \int_{t_1}^\infty t_0 \cdot \text{erfc} \left( \sqrt{\frac{t_0}{4}} \right) dt_0 \right| \leq \alpha_n(t) \cdot \phi_n(\lambda), n = 1, 2, \dots, N+2 \quad (222)$$

These conditions are satisfied with  $\alpha_n(t) = n \cdot 4^{n+1}$  and  $\phi_n(\lambda) = 1$  because, by using the well known upper bound to the complementary error function  $\text{erfc}(x) \leq e^{-x^2}$ ,  $x \geq 0$ , it is possible to prove that:

$$\left| \int_{\lambda t}^\infty dt_{n-1} \cdots \int_{t_1}^\infty t_0 \cdot \text{erfc} \left( \sqrt{\frac{t_0}{4}} \right) dt_0 \right| \leq (\lambda t + 4n) \cdot 4^n \cdot e^{-\lambda t/4} \leq n \cdot 4^{n+1}, n = 1, 2, \dots, N+2 \quad (223)$$

This Theorem also requires that the functions  $p_{d_{ij}^2}^{(n)}(d_{ij}^2)$ ,  $n = 0, 1, \dots, N+1$ , are continuous on  $[0, \infty)$ ,  $p_{d_{ij}^2}^{(N+2)}(d_{ij}^2)$  is piecewise continuous on  $[0, \infty)$  and <sup>12</sup>

$$\int_0^\infty \alpha_n(d_{ij}^2) p_{d_{ij}^2}^{(n)}(d_{ij}^2) dd_{ij}^2 < \infty \quad (224)$$

These conditions are satisfied because, by assumption, the functions  $p_{d_{ij}^2}^{(n)}(d_{ij}^2)$ ,  $n = 0, 1, \dots$ , are continuous and integrable on  $[0, \infty)$ .

Consequently, the asymptotic expansions as  $\text{snr} \rightarrow \infty$  of the integrals that compose the lower bound and the upper bound to the average value of the MMSE are given by:

$$\sum_{n=0}^N \frac{1}{\text{snr}^{n+1}} \cdot k_{\text{LB}_{n+1}} \cdot p_{d_{ij}^2}^{(n)}(0) + \mathcal{O} \left( \frac{1}{\text{snr}^{N+2}} \right) \quad (225)$$

$$\sum_{n=0}^N \frac{1}{\text{snr}^{n+1}} \cdot k_{\text{UB}_{n+1}} \cdot p_{d_{ij}^2}^{(n)}(0) + \mathcal{O} \left( \frac{1}{\text{snr}^{N+2}} \right) \quad (226)$$

<sup>12</sup>The additional requirement that  $\int_0^\infty \alpha_n(d_{ij}^2) p_{d_{ij}^2}^{(n)}(d_{ij}^2) dd_{ij}^2 < \infty$  is due to the fact that the integration interval is semi-infinite.

where the underlying auxiliary asymptotic sequence is  $\{(1/\text{snr})^n, n = 1, 2, \dots\}$  and  $k_{\text{LB}_n}$  and  $k_{\text{UB}_n}$  are given by:

$$k_{\text{LB}_n} = \int_0^\infty dt_{n-1} \int_{t_{n-1}}^\infty dt_{n-2} \cdots \int_{t_1}^\infty \frac{1}{8M(M-1)} \cdot t_0 \cdot \text{erfc}\left(\sqrt{\frac{t_0}{4}}\right) dt_0 \quad (227)$$

$$k_{\text{UB}_n} = \int_0^\infty dt_{n-1} \int_{t_{n-1}}^\infty dt_{n-2} \cdots \int_{t_1}^\infty \frac{1}{2M} \cdot t_0 \cdot \text{erfc}\left(\sqrt{\frac{t_0}{4}}\right) dt_0 \quad (228)$$

We now establish the asymptotic expansions as  $\text{snr} \rightarrow \infty$  of the lower and the upper bounds to the average value of the MMSE in (220) and (221), by using [29, Theorem 1.7.1]. The asymptotic expansion with respect to the asymptotic sequence  $\{\text{snr}^{-n}\}$  as  $\text{snr} \rightarrow \infty$  of the sum of the integrals in (220) and (221) is equal to the sum of the asymptotic expansions with respect to the asymptotic sequence  $\{\text{snr}^{-n}\}$  as  $\text{snr} \rightarrow \infty$  of the individual integrals in (220) and (221), because (225) and (226) are asymptotic expansions of Poincaré type.

Consequently, the asymptotic expansions as  $\text{snr} \rightarrow \infty$  of the lower bound and the upper bound to the average value of the MMSE are given by:

$$\overline{\text{mmse}}_{\text{LB}}(\text{snr}) = \sum_{n=0}^N \frac{1}{\text{snr}^{n+2}} \cdot k_{\text{LB}_{n+1}} \cdot \left( \sum_{i=1}^M \sum_{\substack{j=1 \\ i \neq j}}^M p_{\mathbf{d}_{ij}^2}^{(n)}(0) \right) + \mathcal{O}\left(\frac{1}{\text{snr}^{N+3}}\right) \quad (229)$$

$$\overline{\text{mmse}}_{\text{UB}}(\text{snr}) = \sum_{n=0}^N \frac{1}{\text{snr}^{n+2}} \cdot k_{\text{UB}_{n+1}} \cdot \left( \sum_{i=1}^M \sum_{\substack{j=1 \\ i \neq j}}^M p_{\mathbf{d}_{ij}^2}^{(n)}(0) \right) + \mathcal{O}\left(\frac{1}{\text{snr}^{N+3}}\right) \quad (230)$$

where the underlying auxiliary asymptotic sequence is  $\{(1/\text{snr})^{n+1}, n = 1, 2, \dots\}$  and  $k_{\text{LB}_n}$  and  $k_{\text{UB}_n}$  are also given by (227) and (228), respectively.

The values of  $k_{\text{LB}_n}$  and  $k_{\text{UB}_n}$  can also be computed by using

$$\text{erfc}\left(\sqrt{\frac{t_0}{4}}\right) = \frac{2}{\pi} \int_0^\infty \frac{e^{-\frac{t_0}{4}(x^2+1)}}{x^2+1} dx \quad (231)$$

in (227) and (228), so that:

$$k_{\text{LB}_n} = \frac{2}{\pi} \int_0^\infty \frac{1}{8M(M-1)} \cdot \frac{n \cdot 4^{n+1}}{(x^2+1)^{n+2}} dx = \frac{1}{8M(M-1)} \cdot \frac{n \cdot 4^{n+1}}{\sqrt{\pi}} \cdot \frac{\Gamma(n+3/2)}{\Gamma(n+1/2)} \quad (232)$$

$$k_{\text{UB}_n} = \frac{2}{\pi} \int_0^\infty \frac{1}{2M} \cdot \frac{n \cdot 4^{n+1}}{(x^2+1)^{n+2}} dx = \frac{1}{2M} \cdot \frac{n \cdot 4^{n+1}}{\sqrt{\pi}} \cdot \frac{\Gamma(n+3/2)}{\Gamma(n+1/2)} \quad (233)$$

where  $\Gamma(\cdot)$  is the Gamma function.

Finally, since the functions  $p_{\mathbf{d}_{ij}^2}^{(n)}(\mathbf{d}_{ij}^2)$  are continuous and integrable on  $[0, \infty)$  for arbitrarily large  $n$  and (222) also holds for arbitrarily large  $n$ , it is possible to let  $N$  go to infinity in the asymptotic expansions of the lower and upper bounds to the average value of the MMSE in (229) and (230), respectively.

APPENDIX D  
PROOF OF LEMMA 4

The lower and upper bounds to the average value of the mutual information are obtained from the upper and lower bounds to the average value of the MMSE, respectively, by using (27) as follows:

$$\bar{I}_{\text{LB}}(\text{snr}) = \log M - \int_{\text{snr}}^{\infty} \overline{\text{mmse}}_{\text{UB}}(\xi) d\xi \quad (234)$$

$$\bar{I}_{\text{UB}}(\text{snr}) = \log M - \int_{\text{snr}}^{\infty} \overline{\text{mmse}}_{\text{LB}}(\xi) d\xi \quad (235)$$

We now obtain the asymptotic expansions (as  $\text{snr} \rightarrow \infty$ ) of the lower bound and the upper bound to the average value of the mutual information directly from the asymptotic expansions (as  $\text{snr} \rightarrow \infty$ ) of the upper bound and the lower bound to the average value of the MMSE, respectively, by capitalizing on [29, Theorem 1.7.6].<sup>13</sup> This Theorem requires that:

$$\int_{\text{snr}}^{\infty} \overline{\text{mmse}}_{\text{LB}}(\xi) d\xi < \infty \quad (236)$$

$$\int_{\text{snr}}^{\infty} \overline{\text{mmse}}_{\text{UB}}(\xi) d\xi < \infty \quad (237)$$

and

$$\int_{\text{snr}}^{\infty} \frac{1}{\xi^{n+1}} d\xi < \infty, \quad n = 1, 2, \dots \quad (238)$$

It is immediate to show (e.g. by substituting (220) and (221) in (236) and (237), respectively, and trivially bounding the value of the integrals) that the integrals (236) and (237) exist for  $\text{snr} > 0$  and, likewise, that the integral (238) also exists for  $\text{snr} > 0$ . Consequently, the asymptotic expansion as  $\text{snr} \rightarrow \infty$  of the lower and upper bounds to the average value of the mutual information are given by:

$$\bar{I}_{\text{LB}}(\text{snr}) = \log M - \sum_{n=0}^N \frac{1}{\text{snr}^{n+1}} \cdot k'_{\text{LB}_{n+1}} \cdot \left( \sum_{i=1}^M \sum_{\substack{j=1 \\ i \neq j}}^M p_{\mathbf{d}_{ij}^2}^{(n)}(0) \right) + \mathcal{O}\left(\frac{1}{\text{snr}^{N+2}}\right) \quad (239)$$

$$\bar{I}_{\text{UB}}(\text{snr}) = \log M - \sum_{n=0}^N \frac{1}{\text{snr}^{n+1}} \cdot k'_{\text{UB}_{n+1}} \cdot \left( \sum_{i=1}^M \sum_{\substack{j=1 \\ i \neq j}}^M p_{\mathbf{d}_{ij}^2}^{(n)}(0) \right) + \mathcal{O}\left(\frac{1}{\text{snr}^{N+2}}\right) \quad (240)$$

where the underlying asymptotic sequence is  $\{(1/\text{snr})^n, n = 1, 2, \dots\}$  and  $k'_{\text{LB}_n}$  and  $k'_{\text{UB}_n}$  are given by:

$$k'_{\text{LB}_n} = \frac{1}{n} \cdot k_{\text{UB}_n} = \frac{1}{2M} \cdot \frac{4^{n+1}}{\sqrt{\pi}} \cdot \frac{\Gamma(n+3/2)}{\Gamma(n+1/2)} \quad (241)$$

<sup>13</sup>Note that we determine the asymptotic expansions of the lower and upper bounds to the average mutual information directly from the asymptotic expansions of the upper and lower bounds to the average MMSE in conjunction with (27). Alternatively, we can also determine asymptotic expansions of the lower and upper bounds to the average mutual information by adopting a procedure identical to that in the average MMSE case leveraging the bounds to the mutual information in Lemma 2.



$$k'_{\text{UB}_n} = \frac{1}{n} \cdot k_{\text{LB}_n} = \frac{1}{8M(M-1)} \cdot \frac{4^{n+1}}{\sqrt{\pi}} \cdot \frac{\Gamma(n+3/2)}{\Gamma(n+1/2)} \quad (242)$$

We can also let  $N$  go to infinity in the asymptotic expansions of the lower and upper bounds to the average mutual information in (239) and (240), respectively.

## APPENDIX E

### PROOFS OF THEOREMS 1 AND 2

The proofs capitalize on the asymptotic expansions of the upper and lower bounds to the average value of the MMSE and the average value of the mutual information in Lemmas 3 and 4, respectively, given by:

$$\overline{\text{mmse}}_{\text{LB}}(\text{snr}) = k_{\text{LB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) \cdot \frac{1}{\text{snr}^{d+1}} + \mathcal{O}\left(\frac{1}{\text{snr}^{d+2}}\right) \quad (243)$$

$$\overline{\text{mmse}}_{\text{UB}}(\text{snr}) = k_{\text{UB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) \cdot \frac{1}{\text{snr}^{d+1}} + \mathcal{O}\left(\frac{1}{\text{snr}^{d+2}}\right) \quad (244)$$

and

$$\bar{\text{I}}_{\text{LB}}(\text{snr}) = \log M - k'_{\text{LB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) \cdot \frac{1}{\text{snr}^d} + \mathcal{O}\left(\frac{1}{\text{snr}^{d+1}}\right) \quad (245)$$

$$\bar{\text{I}}_{\text{UB}}(\text{snr}) = \log M - k'_{\text{UB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) \cdot \frac{1}{\text{snr}^d} + \mathcal{O}\left(\frac{1}{\text{snr}^{d+1}}\right) \quad (246)$$

Let us define the functions:

$$f(\text{snr}) = \text{snr}^{d+1} \cdot \overline{\text{mmse}}(\text{snr}) \quad (247)$$

and

$$g(\text{snr}) = \text{snr}^d \cdot (\log M - \bar{\text{I}}(\text{snr})) \quad (248)$$

as well as the functions:

$$f_1(\text{snr}) = \begin{cases} f(\text{snr}), & k_{\text{LB}_d} \cdot \sum_{\substack{i,j=1 \\ i \neq j}}^M p_{d_{ij}^2}^{(d-1)}(0) \leq f(\text{snr}) \leq k_{\text{UB}_d} \cdot \sum_{\substack{i,j=1 \\ i \neq j}}^M p_{d_{ij}^2}^{(d-1)}(0) \\ k_{\text{LB}_d} \cdot \sum_{\substack{i,j=1 \\ i \neq j}}^M p_{d_{ij}^2}^{(d-1)}(0), & f(\text{snr}) \leq k_{\text{LB}_d} \cdot \sum_{\substack{i,j=1 \\ i \neq j}}^M p_{d_{ij}^2}^{(d-1)}(0) \\ k_{\text{UB}_d} \cdot \sum_{\substack{i,j=1 \\ i \neq j}}^M p_{d_{ij}^2}^{(d-1)}(0), & f(\text{snr}) \geq k_{\text{UB}_d} \cdot \sum_{\substack{i,j=1 \\ i \neq j}}^M p_{d_{ij}^2}^{(d-1)}(0) \end{cases} \quad (249)$$

$$f_2(\text{snr}) = \begin{cases} f(\text{snr}) - k_{\text{UB}_d} \cdot \sum_{\substack{i,j=1 \\ i \neq j}}^M p_{d_{ij}^2}^{(d-1)}(0), & f(\text{snr}) \geq k_{\text{UB}_d} \cdot \sum_{\substack{i,j=1 \\ i \neq j}}^M p_{d_{ij}^2}^{(d-1)}(0) \\ 0, & \text{otherwise} \end{cases} \quad (250)$$

$$f_3(\text{snr}) = \begin{cases} -f(\text{snr}) + k_{\text{LB}_d} \cdot \sum_{\substack{i,j=1 \\ i \neq j}}^M p_{d_{ij}^2}^{(d-1)}(0), & f(\text{snr}) \leq k_{\text{LB}_d} \cdot \sum_{\substack{i,j=1 \\ i \neq j}}^M p_{d_{ij}^2}^{(d-1)}(0) \\ 0, & \text{otherwise} \end{cases} \quad (251)$$

and the functions:

$$g_1(\text{snr}) = \begin{cases} g(\text{snr}), & k'_{\text{UB}_d} \cdot \sum_{i \neq j}^M p_{d_{ij}^2}^{(d-1)}(0) \leq g(\text{snr}) \leq k'_{\text{LB}_d} \cdot \sum_{i \neq j}^M p_{d_{ij}^2}^{(d-1)}(0) \\ k'_{\text{UB}_d} \cdot \sum_{i \neq j}^M p_{d_{ij}^2}^{(d-1)}(0), & g(\text{snr}) \leq k'_{\text{UB}_d} \cdot \sum_{i \neq j}^M p_{d_{ij}^2}^{(d-1)}(0) \\ k'_{\text{LB}_d} \cdot \sum_{i \neq j}^M p_{d_{ij}^2}^{(d-1)}(0), & g(\text{snr}) \geq k'_{\text{LB}_d} \cdot \sum_{i \neq j}^M p_{d_{ij}^2}^{(d-1)}(0) \end{cases} \quad (252)$$

$$g_2(\text{snr}) = \begin{cases} g(\text{snr}) - k'_{\text{LB}_d} \cdot \sum_{i \neq j}^M p_{d_{ij}^2}^{(d-1)}(0), & g(\text{snr}) \geq k'_{\text{LB}_d} \cdot \sum_{i \neq j}^M p_{d_{ij}^2}^{(d-1)}(0) \\ 0, & \text{otherwise} \end{cases} \quad (253)$$

and

$$g_3(\text{snr}) = \begin{cases} -g(\text{snr}) + k'_{\text{UB}_d} \cdot \sum_{i \neq j}^M p_{d_{ij}^2}^{(d-1)}(0), & g(\text{snr}) \leq k'_{\text{UB}_d} \cdot \sum_{i \neq j}^M p_{d_{ij}^2}^{(d-1)}(0) \\ 0, & \text{otherwise} \end{cases} \quad (254)$$

In view of the asymptotic expansions in (243), (244), (245) and (246), it is possible to bound  $f(\text{snr})$  and  $g(\text{snr})$  for  $\text{snr} > \text{snr}_0$ , where  $\text{snr}_0$  is a sufficiently high value of  $\text{snr}$ , as follows:

$$k_{\text{LB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) - c_{\text{LB}} \cdot \frac{1}{\text{snr}} \leq f(\text{snr}) \leq k_{\text{UB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) + c_{\text{UB}} \cdot \frac{1}{\text{snr}} \quad (255)$$

and

$$k'_{\text{UB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) - c'_{\text{LB}} \cdot \frac{1}{\text{snr}} \leq g(\text{snr}) \leq k'_{\text{LB}_d} \cdot \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(d-1)}(0) + c'_{\text{UB}} \cdot \frac{1}{\text{snr}} \quad (256)$$

where  $c_{\text{LB}}$ ,  $c_{\text{UB}}$ ,  $c'_{\text{LB}}$  and  $c'_{\text{UB}}$  are positive constants. Note thus that  $f_2(\text{snr})$ ,  $f_3(\text{snr})$ ,  $g_2(\text{snr})$  and  $g_3(\text{snr})$  can be bounded as follows:

$$|f_2(\text{snr})| = f_2(\text{snr}) \leq c_{\text{UB}} \cdot \frac{1}{\text{snr}} = \mathcal{O}\left(\frac{1}{\text{snr}}\right) \quad (257)$$

$$|f_3(\text{snr})| = f_3(\text{snr}) \leq c_{\text{LB}} \cdot \frac{1}{\text{snr}} = \mathcal{O}\left(\frac{1}{\text{snr}}\right) \quad (258)$$

$$|g_2(\text{snr})| = g_2(\text{snr}) \leq c'_{\text{UB}} \cdot \frac{1}{\text{snr}} = \mathcal{O}\left(\frac{1}{\text{snr}}\right) \quad (259)$$

$$|g_3(\text{snr})| = g_3(\text{snr}) \leq c'_{\text{LB}} \cdot \frac{1}{\text{snr}} = \mathcal{O}\left(\frac{1}{\text{snr}}\right) \quad (260)$$

Note also that  $f_1(\text{snr})$  and  $g_1(\text{snr})$  are piecewise infinitely differentiable because  $f(\text{snr})$  and  $g(\text{snr})$  are infinitely differentiable and

$$k_{\text{LB}_d} \cdot \sum_{\substack{i,j=1 \\ i \neq j}}^M p_{d_{ij}^2}^{(d-1)}(0) \leq f_1(\text{snr}) \leq k_{\text{UB}_d} \cdot \sum_{\substack{i,j=1 \\ i \neq j}}^M p_{d_{ij}^2}^{(d-1)}(0) \quad (261)$$

and

$$k'_{\text{UB}_d} \cdot \sum_{\substack{i,j=1 \\ i \neq j}}^M p_{d_{ij}^2}^{(d-1)}(0) \leq g_1(\text{snr}) \leq k'_{\text{LB}_d} \cdot \sum_{\substack{i,j=1 \\ i \neq j}}^M p_{d_{ij}^2}^{(d-1)}(0) \quad (262)$$

Therefore, Theorems 1 and 2 follow immediately by noting that:

$$f(\text{snr}) = f_1(\text{snr}) + f_2(\text{snr}) - f_3(\text{snr}) = f_1(\text{snr}) + \mathcal{O}\left(\frac{1}{\text{snr}}\right) \quad (263)$$

and

$$g(\text{snr}) = g_1(\text{snr}) + g_2(\text{snr}) - g_3(\text{snr}) = g_1(\text{snr}) + \mathcal{O}\left(\frac{1}{\text{snr}}\right) \quad (264)$$

and, in view of (263) and (264) together with (44), (45), (46) and (47),

$$\limsup_{\text{snr} \rightarrow \infty} f(\text{snr}) = \limsup_{\text{snr} \rightarrow \infty} f_1(\text{snr}) = \bar{\epsilon}_d \text{ and } \liminf_{\text{snr} \rightarrow \infty} f(\text{snr}) = \liminf_{\text{snr} \rightarrow \infty} f_1(\text{snr}) = \underline{\epsilon}_d \quad (265)$$

and

$$\limsup_{\text{snr} \rightarrow \infty} g(\text{snr}) = \limsup_{\text{snr} \rightarrow \infty} g_1(\text{snr}) = \bar{\epsilon}'_d \text{ and } \liminf_{\text{snr} \rightarrow \infty} g(\text{snr}) = \liminf_{\text{snr} \rightarrow \infty} g_1(\text{snr}) = \underline{\epsilon}'_d \quad (266)$$

## APPENDIX F

### PROOF OF LEMMA 5

We obtain the lower bound to the error probability by using a genie that supplies the receiver with the true input vector and any of the other input vectors with equal probability. Then,

$$\begin{aligned} P_e(\text{snr}; \mathbf{H}) &= \Pr(e|\mathbf{H}) = \frac{1}{M} \sum_{i=1}^M \Pr(e|\mathbf{x} = \mathbf{x}_i, \mathbf{H}) \\ &\geq \frac{1}{M(M-1)} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \Pr(\mathbf{x}_i \rightarrow \mathbf{x}_j | \mathbf{x} = \mathbf{x}_i, \mathbf{H}) \\ &= \frac{1}{M(M-1)} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \frac{1}{2} \cdot \text{erfc} \left( \sqrt{\frac{d_{ij}^2(\mathbf{H}) \text{snr}}{4}} \right) \triangleq P_{\text{eLB}}(\text{snr}; (\mathbf{H})) \end{aligned} \quad (267)$$

where  $\Pr(e|\mathbf{x} = \mathbf{x}_i, \mathbf{H})$  is the probability of error when  $\mathbf{x}_i$  is transmitted given the channel matrix and  $\Pr(\mathbf{x}_i \rightarrow \mathbf{x}_j | \mathbf{x} = \mathbf{x}_i, \mathbf{H})$  is the probability of choosing  $\mathbf{x}_j$  over  $\mathbf{x}_i$  when  $\mathbf{x}_i$  is transmitted given the channel matrix. The inequality is due to the use of a genie.

We obtain the upper bound to the error probability by using the well-known union bound [31]. Then,

$$\begin{aligned} P_e(\text{snr}; \mathbf{H}) &= \Pr(e|\mathbf{H}) = \frac{1}{M} \sum_{i=1}^M \Pr(e|\mathbf{x} = \mathbf{x}_i, \mathbf{H}) \\ &\leq \frac{1}{M} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \Pr(\mathbf{x}_i \rightarrow \mathbf{x}_j | \mathbf{x} = \mathbf{x}_i, \mathbf{H}) \\ &= \frac{1}{M} \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \frac{1}{2} \cdot \text{erfc} \left( \sqrt{\frac{d_{ij}^2(\mathbf{H}) \text{snr}}{4}} \right) \triangleq P_{\text{eUB}}(\text{snr}; \mathbf{H}) \end{aligned} \quad (268)$$

where  $\Pr(e|\mathbf{x} = \mathbf{x}_i, \mathbf{H})$  is the probability of error when  $\mathbf{x}_i$  is transmitted given the channel matrix and  $\Pr(\mathbf{x}_i \rightarrow \mathbf{x}_j | \mathbf{x} = \mathbf{x}_i, \mathbf{H})$  is the probability of choosing  $\mathbf{x}_j$  over  $\mathbf{x}_i$  when  $\mathbf{x}_i$  is transmitted given the channel matrix. The inequality is due to the union bound.

## APPENDIX G PROOF OF LEMMA 6

We follow the previous procedure to determine the asymptotic expansions of the lower and upper bounds to the average error probability. We obtain the lower and upper bounds to the average value of the error probability from the lower and upper bounds to the error probability by averaging over the fading statistics. The bounds are given by:

$$\bar{P}_{\text{eLB}}(\text{snr}) = \mathbb{E}_{\mathbf{H}} \{P_{\text{eLB}}(\text{snr}; \mathbf{H})\} = \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \int_0^\infty \frac{1}{2M(M-1)} \cdot \text{erfc} \left( \sqrt{\frac{d_{ij}^2 \text{snr}}{4}} \right) \cdot p_{d_{ij}^2}(d_{ij}^2) dd_{ij}^2 \quad (269)$$

$$\bar{P}_{\text{eUB}}(\text{snr}) = \mathbb{E}_{\mathbf{H}} \{P_{\text{eUB}}(\text{snr}; \mathbf{H})\} = \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M \int_0^\infty \frac{1}{2M} \cdot \text{erfc} \left( \sqrt{\frac{d_{ij}^2 \text{snr}}{4}} \right) \cdot p_{d_{ij}^2}(d_{ij}^2) dd_{ij}^2 \quad (270)$$

We obtain the asymptotic expansions as  $\text{snr} \rightarrow \infty$  of the lower and upper bounds to the average value of the error probability by capitalizing on [29, Theorem 3.2] and [29, Theorem 1.7.1]. The asymptotic expansions are given by:

$$\bar{P}_{\text{eLB}}(\text{snr}) = \sum_{n=0}^N \frac{1}{\text{snr}^{n+1}} \cdot k''_{\text{LB}_{n+1}} \cdot \left( \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(n)}(0) \right) + \mathcal{O} \left( \frac{1}{\text{snr}^{N+2}} \right) \quad (271)$$

$$\bar{P}_{\text{eUB}}(\text{snr}) = \sum_{n=0}^N \frac{1}{\text{snr}^{n+1}} \cdot k''_{\text{UB}_{n+1}} \cdot \left( \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_{d_{ij}^2}^{(n)}(0) \right) + \mathcal{O} \left( \frac{1}{\text{snr}^{N+2}} \right) \quad (272)$$

where the underlying auxiliary asymptotic sequence is  $\{(1/\text{snr})^n, n = 1, 2, \dots\}$  and  $k''_{\text{LB}_n}$  and  $k''_{\text{UB}_n}$  are given by:

$$k''_{\text{LB}_n} = \int_0^\infty dt_{n-1} \int_{t_{n-1}}^\infty dt_{n-2} \cdots \int_{t_1}^\infty \frac{1}{2M(M-1)} \cdot \text{erfc} \left( \sqrt{\frac{t_0}{4}} \right) dt_0 \quad (273)$$

$$k''_{\text{UB}_n} = \int_0^\infty dt_{n-1} \int_{t_{n-1}}^\infty dt_{n-2} \cdots \int_{t_1}^\infty \frac{1}{2M} \cdot \text{erfc} \left( \sqrt{\frac{t_0}{4}} \right) dt_0 \quad (274)$$

The values of  $k''_{\text{LB}_n}$  and  $k''_{\text{UB}_n}$  can also be computed by using

$$\text{erfc} \left( \sqrt{\frac{t_0}{4}} \right) = \frac{2}{\pi} \int_0^\infty \frac{e^{-\frac{t_0}{4}(x^2+1)}}{x^2+1} dx \quad (275)$$

in (273) and (274), so that:

$$k''_{\text{LB}_n} = \frac{1}{\pi} \int_0^\infty \frac{1}{M(M-1)} \cdot \frac{4^n}{(x^2+1)^{n+1}} dx = \frac{1}{2M(M-1)} \cdot \frac{4^n}{\sqrt{\pi}} \cdot \frac{\Gamma(n+1/2)}{\Gamma(n+1)} \quad (276)$$

$$k''_{\text{UB}_n} = \frac{1}{\pi} \int_0^\infty \frac{1}{M} \cdot \frac{4^n}{(x^2+1)^{n+1}} dx = \frac{1}{2M} \cdot \frac{4^n}{\sqrt{\pi}} \cdot \frac{\Gamma(n+1/2)}{\Gamma(n+1)} \quad (277)$$

where  $\Gamma(\cdot)$  is the Gamma function. These expansions hold, once again, for arbitrarily large  $N$ .

It is also simple to verify the conditions that justify the application of [29, Theorem 3.2] and [29, Theorem 1.7.1]. For the asymptotic expansion of the individual integrals in (269) and (270), it can be shown that

$$\left| \int_{\lambda t}^{\infty} dt_{n-1} \int_{t_{n-1}}^{\infty} dt_{n-2} \cdots \int_{t_1}^{\infty} \operatorname{erfc} \left( \sqrt{\frac{t_0}{4}} \right) dt_0 \right| \leq \alpha_n(t) \cdot \phi_n(\lambda), \quad n = 1, 2, \dots \quad (278)$$

with  $\alpha_n(t) = 4^n$  and  $\phi_n(\lambda) = 1$ . Furthermore, by assumption, the functions  $p_{d_{ij}^2}^{(n)}(d_{ij}^2)$ ,  $n = 0, 1, \dots$ , are continuous and integrable on  $[0, \infty)$  and

$$\int_0^{\infty} \alpha_n(d_{ij}^2) p_{d_{ij}^2}^{(n)}(d_{ij}^2) dd_{ij}^2 < \infty \quad (279)$$

For the asymptotic expansion of the lower and upper bounds to the average value of the error probability, we use the fact that the asymptotic expansions are of Poincaré type to write the asymptotic expansion of the sum as the sum of the asymptotic expansions.

#### ACKNOWLEDGEMENTS

The author acknowledges fruitful discussions with Sergio Verdú.

#### REFERENCES

- [1] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.
- [2] G. J. Foschini and M. J. Gans, "On the limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 315–335, March 1998.
- [3] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, November/December 1999.
- [4] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Transactions on Information Theory – Special Issue on Shannon Theory: Perspective, Trends and Applications*, vol. 48, no. 6, pp. 1319–1343, June 2002.
- [5] A. M. Tulino, A. Lozano, and S. Verdú, "Multiantenna capacity in the low-power regime," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2527–2544, October 2003.
- [6] S. S. (Shitz) and S. Verdú, "The effect of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1302–1327, May 2001.
- [7] A. Lozano, A. M. Tulino, and S. Verdú, "High-SNR power offset in multiantenna communication," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4134–4151, December 2005.
- [8] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, Now Publishers, vol. 1, no. 1, pp. 1–182, 2004.
- [9] A. M. Tulino, A. Lozano, and S. Verdú, "Impact of antenna correlation on the capacity of multiantenna channels," *IEEE Transactions on Information Theory*, vol. 51, no. 7, pp. 2491–2509, July 2005.
- [10] A. Lozano, A. Tulino, , and S. Verdú, "Capacity-achieving input covariance for single-user multi-antenna channels," *IEEE Transactions on Wireless Communications*, vol. 52, no. 3, pp. 662–671, March 2006.
- [11] S. Venkatesan, S. H. Simon, and R. A. Valenzuela, "Capacity of a Gaussian mimo channel with nonzero mean," in *IEEE Vehicular Technology Conference*, October 2003.
- [12] Y.-H. Kim and A. Lapidoth, "On the log determinant of noncentral Wishart matrices," in *IEEE International Symposium on Information Theory*, June-July 2003.
- [13] A. Lozano, A. M. Tulino, and S. Verdú, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," *IEEE Transactions on Information Theory*, vol. 52, no. 7, pp. 3033–3051, July 2006.
- [14] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1261–1282, April 2005.

- [15] D. P. Palomar and S. Verdú, "Gradient of mutual information in linear vector Gaussian channels," *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 141–154, January 2006.
- [16] F. Pérez-Cruz, M. R. D. Rodrigues, and S. Verdú, "MIMO Gaussian channels with arbitrary inputs: Optimal precoding and power allocation," *IEEE Transactions on Information Theory*, vol. 56, no. 3, pp. 1070–1084, March 2010.
- [17] M. Payaró and D. P. Palomar, "Hessian and concavity of mutual information, differential entropy, and entropy power in linear vector Gaussian channels," *IEEE Transactions on Information Theory*, vol. 55, no. 8, pp. 3613–3628, August 2009.
- [18] —, "On optimal precoding in linear vector Gaussian channels with arbitrary input distribution," in *IEEE International Symposium on Information Theory*, June–July 2009.
- [19] M. Lamarca, "Linear precoding for mutual information maximization in MIMO systems," in *International Symposium on Wireless Communications Systems*, September 2009.
- [20] C. Xiao and Y. R. Zheng, "On the mutual information and power allocation for vector Gaussian channels with finite discrete inputs," in *IEEE Global Communications Conference*, November–December 2008.
- [21] C. Xiao, Y. R. Zheng, and Z. Ding, "Globally optimal linear precoders for finite alphabet signals over complex Gaussian channels," *IEEE Transactions on Signal Processing*, vol. 59, pp. 3301–3314, July 2011.
- [22] W. Zeng, C. Xiao, M. Wang, and J. Lu, "On the linear precoder design for MIMO channels with finite-alphabet inputs and statistical CSI," in *IEEE Global Communications Conference*, December 2011.
- [23] M. Wang, C. Xiao, and W. Zeng, "Linear precoding for MIMO multiple access channels with discrete-constellation inputs," in *IEEE Global Communications Conference*, December 2011.
- [24] A. Lozano, A. M. Tulino, and S. Verdú, "Optimum power allocation for multiuser OFDM with arbitrary signal constellations," *IEEE Transactions on Communications*, vol. 56, no. 5, pp. 828–837, May 2008.
- [25] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effects on the capacity of multielement antenna systems," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 502–513, March 2000.
- [26] D. Chizhik, F. R. Farrokhi, J. Ling, and A. Lozano, "Effect of antenna separation on the capacity of blast in correlated channels," *IEEE Communications Letters*, vol. 4, no. 11, pp. 337–339, November 2000.
- [27] P. Driessen and G. J. Foschini, "On the capacity formula for multiple-input multiple-output channels: A geometric interpretation," *IEEE Transactions on Communications*, vol. 47, no. 2, pp. 173–176, February 1999.
- [28] F. Rashid-Farrokhi, G. J. Foschini, A. Lozano, and R. A. Valenzuela, "Link-optimal space-time processing with multiple transmit and receive antennas," *IEEE Communications Letters*, vol. 5, no. 3, pp. 85–87, March 2001.
- [29] N. Bleistein and R. A. Handelsman, *Asymptotic Expansions of Integrals*. New York, NY: Dover, 1986.
- [30] Y. Wu and S. Verdú, "MMSE dimension," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 4857–4879, August 2011.
- [31] J. Proakis and M. Salehi, *Digital Communications, 5th ed.* New York, NY: McGraw-Hill, 2008.
- [32] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.
- [33] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*. New York, NY: Springer-Verlag, 1988.
- [34] E. Scheuer, "Reliability of an m-out-of-n system when component failure induces higher failure rates in survivors," *IEEE Transactions on Reliability*, vol. 37, no. 1, pp. 73–74, April 1988.
- [35] D. P. Palomar and Y. Jiang, "MIMO transceiver design via majorization theory," *Foundations and Trends in Communications and Information Theory, Now Publishers*, vol. 3, no. 4-5, pp. 331–551, 2006.
- [36] W. Zeng, C. Xiao, M. Qang, and J. Lu, "Linear precoding for finite-alphabet inputs over MIMO fading channels with statistical CSI," *IEEE Transactions on Signal Processing*, vol. 60, pp. 3134–3148, June 2012.
- [37] D. Hammarwall, M. Bengtsson, and B. Otterston, "Acquiring partial csi for spatially selective transmission by instantaneous channel norm feedback," *IEEE Transactions on Signal Processing*, vol. 56, no. 3, pp. 1188–1204, March 2008.
- [38] E. Björnson, D. Hammarwall, and B. Otterston, "Exploiting quantized channel norm feedback through conditional statistics in arbitrarily correlated mimo systems," *IEEE Transactions on Signal Processing*, vol. 57, no. 10, pp. 4027–4041, October 2009.
- [39] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, March 1998.